

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 19:04:01 ; Search time 3885 Seconds

(without alignments)
2344,808 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962
Sequence: 1 MATDVNSKNIAVQAQKIL.....KPHLOKLCGGINMLDEBNI 188

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 10%
Listing first 100 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h
-Q/cgna2.1/USPTO.spool.h/US10627571/runat.27072005.154720.27503/app.query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -IOOPT=0 -IOPEXT=0
-INITs-bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdt -LIST=100
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEASize=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10627571.@CGN_1.1.3731.@runat.27072005.154720.27503 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DISPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	ID	Description
1	954	99.2	1892	9	AF070671 Homo sapi
2	954	99.2	1921	6	BD156880 Primer fo
3	954	99.2	1921	6	AX877846 Sequence
4	954	99.2	1921	9	AK001931 Homo sapi

5	950	98.8	573	9	CR457137 Homo sapi
6	950	98.8	1729	6	BD156785 Primer fo
7	950	98.8	1729	6	AX877673 Sequence
8	950	98.8	1729	9	AK001850 Homo sapi
9	950	98.8	1814	9	BC005352 Homo sapi
10	950	98.8	1915	9	AF099935 Homo sapi
11	950	98.8	1988	6	C0726075 Sequence
12	950	98.8	2003	9	AF099936 Homo sapi
13	950	98.8	2081	9	BC007014 Homo sapi
14	950	98.8	112626	9	AC035144 Homo sapi
15	950	98.8	156277	9	AC026795 Homo sapi
16	950	98.8	158057	9	AC138612 Homo sapi
17	948	98.5	1943	6	BD135356 Homo sapi
18	944	98.1	2502	9	AK097284 Homo sapi
19	920	95.6	1728	10	BC009090 Mus muscu
20	920	95.6	184327	10	AC120859 Mus muscu
21	918	95.4	237561	2	AC095257 Rattus no
22	916	95.2	816	6	BD149395 Primer fo
23	916	95.2	816	6	AX869333 Sequence
24	880	91.5	2009	5	AJ720906 Gallus ga
25	856.5	89.0	1766	9	AK097884 Homo sapi
26	844.5	87.8	1559	5	CR760636 Xenopus t
27	842.5	87.6	1193	5	BC072904 Xenopus l
28	698	72.6	3986	5	BC052765 Danio rer
29	668	72.6	205949	5	BX649252 Zebrafish
30	666	69.2	1589	5	AJ720336 Gallus ga
31	651	67.7	1248	6	CQ736656 Sequence
32	651	67.7	1986	6	CQ841622 Sequence
33	651	67.7	1986	9	AK123281 Homo sapi
34	651	67.7	188937	2	AC012678 Homo sapi
35	651	67.7	189796	9	AC073964 Homo sapi
36	633	65.8	110000	2	AC115187_1 Continuation (2 of
37	633	65.8	256285	2	AC115505 Rattus no
38	633	65.8	273225	2	AC129440 Rattus no
39	626	65.1	164258	2	AC118938 Mus muscu
40	626	65.1	195547	2	AC107848 Mus muscu
41	607	62.4	1938	5	BC076797 Xenopus l
42	600	62.4	134497	5	BC053167 Danio rer
43	600	62.4	189797	5	CR407586 Danio rer
44	593	61.6	32360	9	BX927313 Zebrafish
45	591	61.4	2209	9	AC005339 Homo sapi
46	590	61.3	561	6	BC073263 Sequence
47	587	61.0	587	6	BD149633 Primer fo
48	587	61.0	587	6	AX869571 Sequence
49	574	59.7	1357	5	BC053238 Danio rer
50	571	59.4	107568	2	AC144398 Rattus no
51	571	59.4	230310	2	AC120079 Rattus no
52	567	58.9	22307	10	BC032139 Mus muscu
53	567	58.9	180238	2	AC073688 Mus muscu
54	567	58.9	222610	10	AC026385 Mus muscu
55	567	58.9	232190	2	AC074167 Mus muscu
56	556	57.8	2175	5	BC061657 Xenopus l
57	556	57.8	2175	5	AF098933 Homo sapi
58	536	54.8	1200	9	BC063014 Homo sapi
59	527	54.8	1268	6	AX078255 Sequence
60	527	54.8	3116	9	HSM804686 Homo sapi
61	527	54.8	154526	2	AL592111 Homo sapi
62	527	54.8	159148	2	AL592424 Human DNA
63	527	54.8	181230	2	AC149856 Papio anu
64	526	54.7	182245	2	AC149842 Papio anu
65	526	54.7	207952	2	AC151020 Callithrix
66	524	54.5	1196	9	AK027120 Homo sapi
67	520	54.0	149954	2	CR457375 Homo sapi
68	519	54.0	149954	2	AC151645 Desmops n
69	517	53.7	208375	2	AC147191 Oryzomys
70	515	53.5	217685	2	AC149677 Bos tauru
71	514	53.4	1433	10	BC079019 Rattus no
72	514	53.4	174234	10	AC117098 Rattus no
73	514	53.1	1198	10	BC055879 Mus muscu
74	511	53.1	148675	10	AC131769 Mus muscu
75	511	53.1	205702	10	AC140190 Mus muscu
76	511	53.1	206230	10	AC084272 Mus muscu
77	511	53.1	206230	10	AC084272 Mus muscu

```

78      503      52.3      1171      9      AF271774      Homo sapi
79      484.5      50.4      2156      3      AK112519      Homo int
80      401      41.7      829      6      CQ580257      Sequence
81      401      41.7      1821      3      AY095033      Drosophila
82      383.5      39.9      4406      6      CQ580256      Sequence
83      383.5      39.9      175118      3      AC010842      Drosophila
84      383.5      39.9      188272      3      AC005639      Drosophila
85      383.5      39.9      295225      3      AE003461      Drosophila
86      358      37.2      69208      2      AC020466      Drosophila
87      222      30.4      714      6      BD146713      Primer to
88      222      30.4      714      6      AX866651      Sequence
89      232      30.4      1602      6      BD160707      Primer fo
90      232      30.4      1602      6      AX884081      Sequence
91      232      30.4      1602      6      AK024161      Homo sapi
92      250      26.0      340      6      AX898564      Sequence
93      250      26.0      340      6      BD034097      Sequence
94      166      17.3      252      6      AX898571      Sequence
95      166      17.3      252      6      BD034104      Sequence
96      112      11.6      301130      1      AE016763      Escherich
97      106      11.0      1329      3      AK116070      Homo int
98      105.5      11.0      303414      1      AE015938      Clostridi
99      104.5      10.9      3408      8      AF378568      Eremothec
c 100      104.5      10.9      110000      8      AE016816_3      Continuation (4 of

```

ALIGNMENTS

```

RESULT 1      AF070671      1892 bp      mRNA      linear      PRI 21-JUN-1999
LOCUS      AF070671
DEFINITION      Homo sapiens TNF-induced protein GSG-1 mRNA, complete cds.
ACCESSION      AF070671
VERSION      AF070671.1 GI:3978237
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE      1 (bases 1 to 1892)
AUTHORS      Horrevorts,A.J., Fontijn,R.D., van Zonneveld,A.J., de Vries,C.J.,
              ten Cate,J.W. and Pannekoek,H.
              Vascular endothelial genes that are responsive to tumor necrosis
              factor-alpha in vitro are expressed in atherosclerotic lesions,
              including inhibitor of apoptosis protein-1, stannin, and two novel
              genes
JOURNAL      Blood 93 (10), 3418-3431 (1999)
MEDLINE      99252096
PUBMED      10233894
REFERENCE      2 (bases 1 to 1892)
AUTHORS      Horrevorts,A.J.G., Fontijn,R.D. and Pannekoek,H.
TITLE      Direct Submission
JOURNAL      Submitted (05-JUN-1998) Biochemistry, Academic Medical Center,
              Weibergdreef 15, Amsterdam 1105 AZ, The Netherlands
FEATURES
    source
        1..1892
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /cell_type="endothelial"
            /tissue_type="umbilical vein"
            /cistron="GSG-1"
            /codon_start=1
            /product="TNF-induced protein GSG-1"
            /protein_id="AAC83229.1"
            /db_xref="GI:3978238"
            /translation="MATDVENKSLAVQAKILIGKWSKSIATTLIDTSSVDEL
            VVTFREYTONKEAEKRIKILIKTVIKLILYNNQFNODELAKMKFKKVKYQLAMT
            VVSPHOVDTTPKRNVLRLNCEKREMLHOIQRHLTKASHGRVNVFDPHSDCEFLAA
            LNPENFRPHLOKLCDDGINKMLDEENI"

```

Alignment Scores:

Pred. No.:

9.11e-76

Length:

1892

```

Score:      954.00      Matches:      187
Percent Similarity:      99.47%
Conservative:      0
Best Local Similarity:      99.47%
Query Match:      99.17%
DB:      9      Gaps:      0

```

US-10-627-571-2 (1-188) x AF070671 (1-1892)

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
DB      98 ATGGCCACAGATGTCCTTAAATTCACAAAACCGGCCGTCACAGACAAAAGAGATCTTG 157
QY      21 GlyMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB      158 GGTAAATAGTGTCCAAATTCATCCGCCACCTTAAATGAGACAAAGTAGTAGAGTG 217
QY      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnIysIysGlnIaGluIys 60
DB      218 CTGATGAGCTCTTACAGAGTGACACGAGGAGTACACCCAAAACAGAGAGGACAGAG 277
QY      61 IysIleIysAsnLeuIleIleThrValIleIysLeuAlaIleLeuTyArgAsnGln 80
DB      278 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTTTATAGAAATATCAG 337
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIysIysValHisGlnLeuAla 100
DB      338 TTTAATCAAGATGAGCTAGCTGATTCATGAGAAATTTAAGAAAGATTCATCAGCTTGT 397
QY      101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValIleSerArg 120
DB      398 ATGACCGTGTGATGATTCATCAGGTCGATTCATCTTGGACCGGATGTGATTCAGAG 457
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisIleuThrAlaIys 140
DB      458 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCACCGCCACTGCTGCAAG 517
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
DB      518 TCACATGACGCGGTTAATATGTCCTTGAATCTTTTCAGATTTGCAATTTTGGCTGCG 577
QY      161 LeuTyAsnProPheGlyAsnIleIysProHisIleGlnIysLeuIysAspGlyIleAsn 180
DB      578 TGTATTAATCCCTTTGGGAATTTTAAACCCCACTTACAAAACATAATGATGATGATCAAC 637
QY      181 IysMetLeuAspGluGluAsnIle 188
DB      638 AAAATGTTGATGAAGAGAAACATA 661

```

```

RESULT 2      BD156880      1921 bp      DNA      linear      PAT 17-JAN-2003
LOCUS      BD156880
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD156880
VERSION      BD156880.1 GI:27862638
KEYWORDS      JP 2002191363-A/11723.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE      1 (bases 1 to 1921)
AUTHORS      Oca,T., Isogai,T., Nishikawa,T., Haysashi,K., Saito,K., Yamamoto,J.,
              Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
              Primer for synthesizing full-length cDNA and use thereof
              Patent: JP 2002191363-A 11723 09-JUL-2002;
JOURNAL      HELIX RESEARCH INSTITUTE
COMMENT      OS Homo sapiens (human)
PN JP 2002191363-A/11723
PF 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI

```

PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(140) . (703) .

FEATURES
source
1. .1921
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 9.25e-76 Length: 1921
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x BD156880 (1-1921)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAgiLysIleLeu 20
DB 140 ATGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTCAGGCAAAAAGAGATCTTG 199
QY 21 GlyMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 200 GGTAAATGGTGTCCAAATCCATCGCCACCACTTAAATACACAGATAGTAGTG 259
QY 41 LeuAspGluLeuTyraArgValThrArgGluTyThrGlnAsnLysGluAgiLys 60
DB 260 CTGGATGAGCTCTACAGAGTACCGAGGAGTACCCAAAACAAAGAGAGCGAGAG 319
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnGln 80
DB 320 ATCATCAAGAACCTCATCAAGACAGATCATCAAGCTGCGCATTTCTTAAGAAATATCAG 379
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 380 TTTAAATCAAGATAGCTAGATTCATGATGAGAAATTTAAGAGAAAGTTTATCAGCTTGT 439
QY 101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB 440 ATGACCGTGGTCACTTCATCAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 499
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 500 CTGTAAATGAATGACAGAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTGCAG 559
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
DB 560 TCACATGACGACGGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 619
QY 161 LeuTyraAsnProPheGluAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 620 TTGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTATGTATCAAC 679
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 680 AAAATGTTGATGACAGAGAACATTA 703

RESULT 3
AX877846 1921 bp DNA linear PAT 17-DEC-2003
LOCUS AX877846
DEFINITION Sequence 12751 from Patent EP1074617.
ACCESSION AX877846
VERSION AX877846.1 GI:40032582
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1
AUTHORS Oca, T., Isogai, T., Nishikawa, T., Hayaeshi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12751 07-FEB-2001;
Research Association for Biotechnology (JTB)

FEATURES
source
1. .1921
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS

140. .706
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89903.1"
/db_xref="GI:40032583"
/translations="MATDVPSNKILAVQAKTIIGKMSKSIATTLIDPSSSELDL
YRTREYTONKKAETIKRLITKIVLKAILIRNPNQDELAMEKFKKHQILAMT
VSEFQVDYTFPBNVLSRLNECREMLHQIIOHLLTAKSHGRVNNVFDHPSDFELAA
LYNPFQNFPHLQKLCGINKMDEBNTI"

ORIGIN

Alignment Scores:
Pred. No.: 9.25e-76 Length: 1921
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x AX877846 (1-1921)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAgiLysIleLeu 20
DB 140 ATGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTCAGGCAAAAAGAGATCTTG 199
QY 21 GlyMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 200 GGTAAATGGTGTCCAAATCCATCGCCACCACTTAAATAGACACAGATAGTAGTG 259
QY 41 LeuAspGluLeuTyraArgValThrArgGluTyThrGlnAsnLysGluAgiLys 60
DB 260 CTGGATGAGCTCTACAGAGTACCGAGGAGTACCCAAAACAAAGAGAGCGAGAG 319
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnGln 80
DB 320 ATCATCAAGAACCTCATCAAGACAGATCATCAAGCTGCGCATTTCTTAAGAAATATCAG 379
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 380 TTTAAATCAAGATAGCTAGATTCATGATGAGAAATTTAAGAGAAAGTTTATCAGCTTGT 439
QY 101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB 440 ATGACCGTGGTCACTTCATCAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 499
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 500 CTGTAAATGAATGACAGAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTGCAG 559
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
DB 560 TCACATGACGACGGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 619
QY 161 LeuTyraAsnProPheGluAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 620 TTGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTATGTATCAAC 679
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 680 AAAATGTTGATGACAGAGAACATTA 703

RESULT 4
AK001931 1921 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FL111069 f1s, clone PLACE1004930, highly similar
DEFINITION to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK001931
VERSION AK001931.1 GI:7023502
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Odayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Houchi, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuma, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Nomiyama, H., Satoh, N., Takani, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumaga, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togauchi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
TITLE Complete sequencing and characterization of 21,243 full-length
human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702035
REFERENCE
AUTHORS 2
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Aotsuma, S., Yoshikawa, Y.,
Matsunawa, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Sasaki, N.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE
AUTHORS 3 (bases 1 to 1921)
Isogai, T. and Otsuki, T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@r1.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
1..1921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004930"

ORIGIN
Alignment Scores:
Pred. No.: 9,25e-76 Length: 1921
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: Gaps: 0
US-10-627-571-2 (1-188) x AK001931 (1-1921)
QY 1 MetAlAThPaSPVaIPheAsnSerLySaNleuAlaValGlnAlGlnLySleleu 20
|||||
140 ATGGCCACAGAGTCTTTAATTCAAAACCTGGCCGTTCAAGCACAAGAGAGTCTTG 199
DB 21 GLYSMeTvaISerLySerIleAlaThrThrLeuIleAspAspThSerSerGluVal 40
|||||
200 GGTAAATGCTGTCCTCAATCCATCCGACCTTATATGACGACGACGATGAGGTG 259
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLysGlnAgiuLys 60
|||||
260 CTGATGACCTCTACAGAGTGACAGGAGGTACACCAAAACAGAGAGGAGAGAG 319
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLysAlaIleLeuTyArgAsnGln 80
|||||
320 ATCATCAAGAACCTCATCAAGACATCAAGCTGCGCATCTTTATAGAAATATCATG 379
DB 81 PheAsnGluAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
|||||
380 TTTAATCAAGATGAGCTACATTCATGAGAAATTTAAGAGAAATGATCATGCTTGT 439
QY 101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValIleuSerArg 120
|||||
440 ATGACCGTGTGATTCATCATGAGTGTATATCTTTCAGCCGGAATGTATTCAGG 499
DB 121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
|||||
500 CTGTAAAGAAATGAGAGAGAGATGCTGCACCAATCATTCACCGCACCTGCAAG 559
QY 141 SerHisGluArgValAsnAsnValPheAspHisPheSerArgCysGluPheLeuAlaIle 160
|||||
560 TACATGACGCGGCTTATATATGTGTTGATCATCTTTTCAGATTGCAATTTTGGCTGCC 619
DB 161 LeuTyAsnProPheGluAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
|||||
620 TTGTATATCTTTTGGGAATTTAAACCCCACTTACAAAACATAATGTATGATCATAC 679
QY 181 LysMetLeuAspGluGluLysIle 188
|||||
680 AAAATGTTGATGAAGAGAACATCA 703
DB
RESULT 5
CR457137 573 bp mRNA linear PRI 03-JUN-2004
LOCUS Homo sapiens full open reading frame cDNA clone RZPD034H127D for
DEFINITION gene TNFAPB8, tumor necrosis factor, alpha-induced protein 8;
complete cds, incl. stopcodon.
ACCESSION CR457137
VERSION CR457137.1 GI:48146390
KEYWORDS Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 573)
Ebert, L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.
TITLE Cloning of human full open reading frames in Gateway(TM) system
JOURNAL entry vector (pDONR201)
Unpublished

REFERENCE
2 (pages 1 to 573)
AUTHORS
Ebert, J., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.
TITLE
Submitted (03-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer

REFERENCE	2 (pages 1 to 573)
AUTHORS	Ebert, L., Schlick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.
TITLE	Direct Submission
JOURNAL	Submitted (03-JUN-2004) RZP Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
COMMENT	RZPD; RZPD0834H127D, ORFNO 1910

FEATURES	Location/Qualifiers
Source	1. .573

ORIGIN

Alignment Scores:

Pred. No.:	6.04e-76	Length:	57
Score:	950.00	Matches:	186
Percent Similarity:	99.47%	Conservative:	1
Best Local Similarity:	98.94%	Mismatches:	1
Query Match:	98.75%	Indels:	0
DB:	9	Gaps:	0

US-10-627-571-2 (1-188) x CR457137 (1-573).

[illegible]

Qy	41	LeuAspGluLeuYrArgValThrArgIuYrThrGlnAsnIuYsGluAlaGluIuYs	60
Db	127	CYGAATAGAGCTTACAGAGTGACCGAGAGTACCCAAACAGAGAGGACAGAGAG	186
Qy	61	LysIleYsAsnLeuIleYrThrValIleYsLeuAlaIleLeuYrArgAsnAsnGln	80
Db	187	ATCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCATCTTTATAGGAATATACG	246
Qy	81	PheAsnGlnAspGluLeuAlaLeuMetGluYsPheYsIuYsValHisGlnLeuAla	100
Db	247	TTTATATCAAGATGAGCTACGATTTATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT	306
Qy	101	MetThrValValSerPheHisGlnValAspYrThrPheAspArgAsnValLeuSerArg	120
Db	307	ATGACCGGTGGTCAGTTTCCATCAGAGTGATTTATACCTTTGACCGGAAATGTGTATCCAG	366
Qy	121	LeuLeuAsnGluYsArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaYs	140
Db	367	CTGTTAATATGATCAGAGAGATCTGCAACCAATCATTCAGCGGCACCTCAGCTCCAG	426
Qy	141	SerHisGluYrArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa	160
Db	427	TCAATGACCGGGTTAATATGATGTGGATATTTTCAGATGTGAAATTTTGGCTGCC	486
Qy	161	LeuYrAsnProPheGluYsAsnPheYsProHisLeuGlnIuYsLeuYsAspGluYrIleAsn	180
Db	487	TTGTATATCTCTTTTGGGAATTTTAAACCCACATTACAAAACATATGATGATGATCAAC	546
Qy	181	LysMetLeuAspGluGluAsnIle	188
Db	547	AAATGTGTAGTGAAGAAACATT	570

LOCUS	BD156785	1729 bp	DNA	linear	PAT 17-JAN-2003
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				

ACCESSION BD156785
VERSION BD156785.1 GI:27862543
KEYWORDS JP 2002191363-A/11628.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE
Primer for synthesizing full-length cDNA and use thereof
JOURNAL
Patent: JP 2002191363-A 11628 09-JUL-2002;

COMMENT

PN	JP 2002191363-A/11628
PD	09-JUL-2002
PF	28-JUL-2000 JP 2000280990
PI	TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI	SAITO,
PI	KUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU
PI	KEIICHI NAGAI, TETSUJI OTSUKI
PC	C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/09, C12O1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
FT	Primer for synthesizing full-length cDNA and use thereof FH Key
CDS	Location/Qualifiers (71). (1664).

ORIGIN	Alignment Scores:
Pred. No.:	1.89e-75
	Length: 1722

Score: 950.00 Matches: 186
 Percent Similarity: 98.47% Conservative: 1
 Best Local Similarity: 98.94% Mismatches: 1
 Query Match: 98.75% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x BD156785 (1-1729)

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
      :::
DB      101 GTGGCCACGATGTCCTTAATTCCAAAAACCTGGCCGTTCCAGGACCAAAAAGAGATCTTG 160
QY      21 GlyLysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      |||
DB      161 GGTAAATGTCCTCCAAATCCATCCGACCACTTAATAGACACACAGTAGTAGG 220
QY      41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnIysIleGlnAlaGluLys 60
      |||
DB      221 CTGGATGACCTTACAGAGTACACAGGAGTACACCAAAAACAAGAGGACAGAGAAG 280
QY      61 LysIleLysAsnLeuIleIysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
      |||
DB      281 ATCATCAAGAACTCATCAAGACAGATCATCAAGCTGCCATCTTATAGGAATTAATCAG 340
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIleValHisGlnLeuAla 100
      |||
DB      341 TTTAATCAAGATGACCTAGCATTTGATGGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT 400
QY      101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
      |||
DB      401 ATGACCGTGGTCAGTTTCCATCAGTGGATTATACCTTTCGACCGGAATGTTATCCAG 460
QY      121 LeuLeuAsnGlnCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
DB      461 CTGTAAAGAAAGACGAGAGATGCTGCACCAATATATTCAGGCCACCTCAGCTGCAG 520
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
      |||
DB      521 TCACATGAGACGGGTAAATATATGTTGATTCATTTTTCAGATTTGTAATTTGGCTGCC 580
QY      161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnIysLeuCysAspGlyIleAsn 180
      |||
DB      581 TTGTATATATCCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTATGATGATGATCAAC 640
QY      181 LysMetLeuAspGluGluAsnIle 188
      |||
DB      641 AAAATGTTGATGAAGAACAACATA 664

RESULT 7
AX877673      1729 bp      DNA      linear      PAT 17-DEC-2003
LOCUS      AX877673
DEFINITION      Sequence 12578 from Patent EP1074617.
ACCESSION      AX877673
VERSION      AX877673.1
KEYWORDS      GI:40032409
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Ota,T., Isogai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
Primers for synthesising full-length cDNA and their use
Patent: EP 1074617-A 12578 07-FEB-2001;
Research Association for Biotechnology (JBP)
LOCATION/Qualifiers
1. 1729
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
71..667
/note="unnamed protein product"
/codon_start=1

CDS

```

/protein_id="CAE98929.1"
 /db_xref="GI:40032410"
 /translation="MHSRSESKEVATDVNSKNLAVOAKKILGKMVSKSATLLID
 DTSSVLDLKRVTREYTONKKEAKKIKLKITVYIKALILRNQPNODELALMEKF
 KKKVQLMTWVSFHQVDYTRDRNVLSHLMECREMHLQIILQRHLTASHGKVNVPD
 HPSDEFIAALYNPPGNFKPHLOKLCIDGINKMLDEENT"

ORIGIN

Alignment Scores:

Pred. No.: 1,89e-75 Length: 1729
 Score: 950.00 Matches: 186
 Percent Similarity: 99.47% Conservative: 1
 Best Local Similarity: 98.94% Mismatches: 1
 Query Match: 98.75% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x AX877673 (1-1729)

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
      :::
DB      101 GTGGCCACGATGTCCTTAATTCCAAAAACCTGGCCGTTCCAGGACCAAAAAGAGATCTTG 160
QY      21 GlyLysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      |||
DB      161 GGTAAATGTCCTCCAAATCCATCCGACCACTTAATAGACACACAGTAGTAGG 220
QY      41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnIysIleGlnAlaGluLys 60
      |||
DB      221 CTGGATGACCTTACAGAGTACACAGGAGTACACCAAAAACAAGAGGACAGAGAAG 280
QY      61 LysIleLysAsnLeuIleIysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
      |||
DB      281 ATCATCAAGAACTCATCAAGACAGTATCAAGCTGCCATCTTATAGGAATTAATCAG 340
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIleValHisGlnLeuAla 100
      |||
DB      341 TTTAATCAAGATGACCTAGCATTTGATGGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT 400
QY      101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
      |||
DB      401 ATGACCGTGGTCAGTTTCCATCAGGAGTATACCTTTCGACCGGAATGTTATCCAG 460
QY      121 LeuLeuAsnGlnCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
DB      461 CTGTAAAGAAAGACGAGAGATGCTGCACCAATATATTCAGGCCACCTCAGCTGCAG 520
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
      |||
DB      521 TCACATGAGACGGGTAAATATATGTTGATTCATTTTTCAGATTTGTAATTTGGCTGCC 580
QY      161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnIysLeuCysAspGlyIleAsn 180
      |||
DB      581 TTGTATATATCCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTATGATGATGATCAAC 640
QY      181 LysMetLeuAspGluGluAsnIle 188
      |||
DB      641 AAAATGTTGATGAAGAACAACATA 664

RESULT 8
AK001850      1729 bp      mRNA      linear      PRI 30-JAN-2004
LOCUS      AK001850
DEFINITION      Homo sapiens cDNA FLJ10988 fis, clone PLACB1001920, highly similar
to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION      AK001850
VERSION      AK001850.1
KEYWORDS      GI:7023373
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Makamatsu,A., Hayaishi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,

```


JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED
12477932

REFERENCE
2 (bases 1 to 1814)

AUTHORS
Straussberg, R.

TITLE
Direct Submission

JOURNAL
Submitted (27-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
source
1..1814
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12451 IMAGE:3997650"
/tissue_type="Bladder carcinoma"
/clone_id="N1H_MGC_53"
/lab_host="DH108"
/note="vector: PDNR-L1B"
1..1814
/gene="TNFAIP8"
/note="synonyms: G93-1, MDC-3.13, SCC-S2"
/db_xref="LOCUSID:25816"
144..740
/gene="TNFAIP8"
/codon_start=1
/product="TNFAIP8 protein"
/protein_id="AAH05352.1"
/db_xref="GI:13529164"
/db_xref="LOCUSID:25816"
/translation="MHSEAESKEVATDVFNKSLAVQAOKKILGKMSKSIATTLID
DTSSVDELVYVREYTONKKEAEKIIKNLIKIVIKLAILYRNNOFQDELALMEKF
KKVHQLAMTVTSFHOVDYTPDRNVLSRLNLCREMLHQIILRHILTAKSHGRVNVFP
HSDCEFLAALVNPGRFPHQIKICDGINKLMDENT"
180..737
/gene="TNFAIP8"
/note="DUF758; Region: Domain of unknown function
(DUF758). Family of eukaryotic proteins with unknown
function, which are induced by tumour necrosis factor"
/db_xref="CCD:pfam05527"

ORIGIN
Alignment Scores:
Pred. No.: 1..98e-75 Length: 1814
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 96.75% Indels: 0
DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x BC005352 (1-1814)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysLysIleLeu 20
:::|||||
Db 174 GTGGCCACGATGTCCTTAATTCACAAACCTGGCCGCTTCAGGCACAAAGAAAGATCTTG 233

QY 21 GILysMetValSerIysSerIleAlaThrThrIleuIleAspAspThrSerSerGluVal 40
|||||
Db 234 GGTTAAATGATGTCCTTAATTCATCCGCCACCTTTAATAGACACAAAGTATGATG 293

QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnIlyLysGluAlaGluIys 60
|||||
Db 294 CTGACGAGCTCTTACAGAGTGACGAGGAGTACACCAAAACAAAGAGAGGACAGAA 353

QY 61 LysIleIysAsnIleIleIysThrValIleIysLeuAlaIleLeuTyrArgAsnGln 80
|||||
Db 354 ATCATCAAGAACCTCATCAAGACAGCTCATCAAGCTGCATCTTTAATAGAAATATCAG 413

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIlyLysValIleGlnLeuAla 100
|||||
Db 414 TTTAATCAAGATGACTTGCATTTATGAGAAATTTAAGAGAAAGTTATCAGCTTCT 473

QY 101 MetThrValIleSerPheHisGlnValAspTyrThrPheAspArgAsnValIleSerArg 120
|||||
Db 474 ATGACCGTGTCAGTTTCCATCAGGTGATTTATACCTTTGACCGGAATGTGTATCCAG 533

QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIys 140
|||||
Db 534 CTGTTAAATGATGACAGAGATGCTGCACCAATCATTCAGCCGCTCATGCTCCAG 593

QY 141 SerHisGluArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
|||||
Db 594 TCACATGACGCGTTAATATATGCTTTGATCATTTTCAGATTTGGAATTTTGCTGCC 653

QY 161 LeuTyrAsnProPheGluAsnPhelysProHisleuGlnIlyLeuCysAspGlyIleAsn 180
|||||
Db 654 TTGTATTAATCCTTTTGGAAATTTTAAACCCACCTTACAAAAACTATGTGATGTATCAAC 713

QY 181 LysMetLeuAspGluGluAsnIle 188
|||||
Db 714 AAAATGTGATGAGAGAACATA 737

RESULT 10
AF099935
LOCUS Homo sapiens MDC-3.13 isoform 2 mRNA, complete cds. PRI 12-NOV-1998
DEFINITION AF099935
ACCESSION AF099935
VERSION AF099935.1 GI:3860092
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1915)
Nierfeld, W. and Meyerhans, A.F.
Identification of cellular factors involved in the differentiation
of dendritic cells
Unpublished
2 (bases 1 to 1915)
Nierfeld, W. and Meyerhans, A.F.
Direct Submission
Submitted (13-OCT-1998) Max Planck Institute for Molecular
Genetics, Ihmestrasse 73, Berlin 14195, Germany
Location/Qualifiers
1..1915
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="adherent monocytes treated with GM-CSF and
IL-4"
85..681
/note="MDC-3.13/2"
/codon_start=1
/product="MDC-3.13 isoform 2"
/protein_id="AAC72975.1"
/db_xref="GI:3860093"
/translation="MHSEAESKEVATDVFNKSLAVQAOKKILGKMSKSIATTLID
DTSSVDELVYVREYTONKKEAEKIIKNLIKIVIKLAILYRNNOFQDELALMEKF

CDs

ORIGIN
KKKVLHQLAMTVSFEHQVDYTFEDRNVL SRLNCEKEMLHOI IQRHLTAKSHGVNVNF
HFSDCEFLALVYVPGNFKXPHLQKLCGINKMLDEENI"

Alignment Scores:
Pred. No.: 2,1e-75 Length: 1915
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
Gaps: 0

US-10-627-571-2 (1-188) x AF099935 (1-1915)

QY 1 MetCAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
115 GTGGCCACAGATGCTTAAATCCAAAACCTGGCCCTTCAGGACCAAAAGAGATCTTG 174
QY 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
175 GGTAAATGGTGTCCAATCCATCGCCACCACTTAATAGACGACACAGTAGTAGGTG 234
QY 41 LeuAspGluLeuTyrrArgValThrArgLysThrGlnAsnLysLysGluAlaGluLys 60
235 CTGGATGAGCTTACAGAGTACAGGAGGATACCCAAACAAAGAGGAGGAGAG 234
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnAsnGln 80
295 ATCATCAAGAACCTCATCAAGACAGATCATGAGTGGCCATCTTTATAGAAATATACAG 354
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
355 TTTAAATCAAGATGAGTACAGATTCAGAGAGAAATTTAAGAAAGTTCATCAGCTTGTCT 414
QY 101 MetThrValIleSerPheHisGlnValAspTyrrThrPheAspArgAsnValIleuSerArg 120
415 ATGACCGTGTCTGATTCATCAGTGGATTTATACCTTTACCGGATGTGTTATTCAGG 474
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
475 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCAGTCCAG 534
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
535 TCACATGAGAGCGGTAAATATGTTGATTCATTTTCAATTTGTAATTTTGGCTGCC 594
QY 161 LeuTyrrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
595 TTGTATATATCTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGTATGATTCAC 654
QY 181 LysMetLeuAspGluGluAsnIle 188
655 AAAATGTTGATGAGAGAACATTA 678
Db

RESULT 11
CQ726075 1988 bp DNA linear PAT 03-FEB-2004
LOCUS CQ726075
DEFINITION Sequence 12009 from Patent WO02068579.
ACCESSION CQ726075
VERSION CQ726075.1 GI:42288070
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humaneurons or transcripts, for detecting expression and other uses
therof
JOURNAL Patent: WO 02068579-A 12009 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES location/Qualifiers

source 1. .1988
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 2,18e-75 Length: 1988
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
Gaps: 0

US-10-627-571-2 (1-188) x CQ726075 (1-1988)

QY 1 MetCAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
204 GTGGCCACAGATGCTTAAATCCAAAACCTGGCCCTTCAGGACCAAAAGAGATCTTG 263
QY 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
264 GGTAAATGGTGTCCAATCCATCGCCACCACTTAATAGACGACACAGTAGTAGGTG 323
QY 41 LeuAspGluLeuTyrrArgValThrArgLysThrGlnAsnLysLysGluAlaGluLys 60
324 CTGGATGAGCTTACAGAGTACAGGAGGATACCCAAACAAAGAGGAGGAGAGAG 383
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnAsnGln 80
384 ATCATCAAGAACCTCATCAAGACAGATCATGAGTGGCCATCTTTATAGAAATATACAG 443
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
444 TTTAAATCAAGATGAGTACAGATTCAGAGAGAAATTTAAGAAAGTTCATCAGCTTGTCT 503
QY 101 MetThrValIleSerPheHisGlnValAspTyrrThrPheAspArgAsnValIleuSerArg 120
504 ATGACCGTGTCTGATTCATCAGTGGATTTATACCTTTGACCGGATGTGTTATTCAGG 563
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
564 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCAGTCCAG 623
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
624 TCACATGAGAGCGGTAAATATGTTGATTCATTTTCAATTTTGGCTGCC 683
QY 161 LeuTyrrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
684 TTGTATATATCTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGTATGATTCAC 743
QY 181 LysMetLeuAspGluGluAsnIle 188
744 AAAATGTTGATGAGAGAACATTA 767
Db

RESULT 12
AF099936 2003 bp mRNA linear PRI 12-NOV-1998
LOCUS AF099936
DEFINITION Homo sapiens MDC-3.13 isoform 1 mRNA, complete cds.
ACCESSION AF099936
VERSION AF099936.1 GI:3860094
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Niefeld, W. and Meyerhans, A.F.
TITLE Identification of cellular factors involved in the differentiation
of dendritic cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2003)

AUTHORS Niefeld, W. and Meyerhans, A.F.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-1998) Max Planck Institute for Molecular Genetics, Ihmestrasse 73, Berlin 14195, Germany
 FEATURES
 source
 1. 2003
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_type="adherent monocytes treated with GM-CSF and IL-4"
 198. .770
 /note="MDC-3.13/1"

CDS

/codon_start=1
 /product="MDC-3.13 isoform 1"
 /protein_id="AAC72976.1"
 /db_xref="GI:3860395"
 /translation="MAVATDVFNKSLVAQAKILGKRVKSATTLIDTTSSEVLD
 ELRYRREYQNKKEAKIKILIKIVILLYRNQFNOBELAMEKFKKVLQLA
 MIVSPHOVDYTFEDRNVL SRLINCEGEMHQIIRHLTKSHGRVNVNVDHPSDCEFL
 AALVDFGNFKRHLQKLCDCGINKMDEENI"

ORIGIN

Alignment Scores:
 Pred. No.: 2,2e-75 Length: 2003
 Score: 950.00 Matches: 186
 Percent Similarity: 99.47% Conservative: 1
 Best Local Similarity: 98.94% Mismatches: 1
 Query Match: 98.75% Indels: 0
 DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x AF099936 (1-2003)

QY 1 MetAlaThaAspValPhehSenSerLyAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
 Db 204 GTGGCCACGAGATGCTTAATTCACAAAACCTGGCCGTCACGACACAAAAGAGATCTTG 263
 QY 21 GilytMetValSerlySerIleAlaThrThrLeuIlaAspAspThrSerSerGluVal 40
 Db 264 GGTAAATGCTGCCAATCCATCGCCACCACTTAAAGACACACAAAGTAGTGGG 323
 QY 41 LeuAspGluLeuTyrrArgValThrArgGluTyrrThrGlnAsnLysGlnAlaGluLys 60
 Db 324 CTGAGTAGACCTCTACAGAGTAGACAGGAGTAGACCCAAAACAAGAAGGAGGAGAG 383
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnGln 80
 Db 384 ATCATCAAGAACCTCATCAAGACAGCATCAAGCTGCCATCTTAATGAGATTAATCAG 443
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysLysValHisGlnLeuAla 100
 Db 444 TTTAATCAAGATAGCTTAGCATTTGATGAGAAATTTAAGAAAGATTCTCAGCTTGCT 503
 QY 101 MetThrValIaSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
 Db 504 ATGACCGTGATGATTTCCATCAGTGAGATTATACCTTTGACCGAATGTTATATCAG 563
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 564 CTGTTAAAGAAAGCAGAGAGAGATGACCAAAATATTCAGGCCACCTCAGCTGCAAG 623
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
 Db 624 TCACATGACGCGGTTAATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 683
 QY 161 LeuTyrrAsnProhGluLysAsnPhelysProhIleLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 684 TGTATTAATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTATGTATATCAAC 743
 QY 181 LysMetLeuAspGlnGluLysIle 188
 Db 744 AAAATGTTGATGAAGAACAATCA 767

RESULT 13

BC007014

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BC007014 2081 bp mRNA linear PRI 24-FEB-2004
 Homo sapiens tumor necrosis factor, alpha-induced protein 8, mRNA
 (cDNA clone MGC:12346 IMAGE:3930240), complete cds.

BC007014 GI:13937825

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 2081)

Klausner, R.D., Collins, P.S., Wagner, L.H., Derge, J.G.,

Alteich, S.F., Zeeberg, B., Bueow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Uedini, T.B., Toshiyuki, S.,

Carrinzi, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Adams, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richard, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Holly, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahney, J., Helton, E., Kettelman, M., Madan, A., Young, A.C.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Scherchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Kizyanski, M.I., Skalska, U., Smallwood, D.E.,

Schneer, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2081)

Strausberg, R.

Direct Submission

Submitted (30-APR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mdedpax11.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LUNL at: <http://image.llnl.gov>

Series: IRM Plate: 16 Row: D Column: 7

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 7657123.

Location/Qualifiers

1. 2081

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:12346 IMAGE:3930240"

/tissue_type="Brain, primitive neuroectodermal"

/clone_lib="NIH MGC_56"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

1. 2081

FEATURES

source

gene

/gene="TNFAIP8"

/note="synonyms: G32-1, MDC-3.13, SCC-S2"

CDS

/db_xref="LocustID:25816"
237..833
/gene="TNFAIP8"
/codon_start=1
/product="TNFAIP8 protein"
/protein_id="AAH07014.1"
/db_xref="GI:13937826"
/db_xref="LocustID:25816"
/translation="MSEAESEKVAATDVFNKNLAVOQKILGKRVSKIATVTLID
DTSEVLDLRYVREYTONKKEAKIKLITVIVIKALIVNNQFNODELAMEKF
KKVHOLAMTVVSPFHQVDYTFDRLVLSRLNCEGEMHGIQIRHLTRAKSHGVNVED
HPSDCERIALVNPFGNFKPHLQKLCQGIKMDENI"
273..830
/gene="TNFAIP8"
/note="DUF758; Region: Domain of unknown function
(DUF758). Family of eukaryotic proteins with unknown
function, which are induced by tumour necrosis factor."
/db_xref="CDD:pfam05527"

misc_feature

Alignment Scores:
Pred. No.: 2,29e-75 Length: 2081
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x BC007014 (1-2081)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
Db : : : : :
267 GTGGCCACAGATGCTTTTAATCCAAAACCTGGCCCTTACGACACAAAAGAAATCTTG 326
QY 21 GlyLysMetValSerLysSerIleAlaThrThreLeuIleAspAspThrSerSerGluVal 40
Db : : : : :
327 GGTAAATGGTGTCCAAATCCATCGCCACACCTTAAATACACACAAAGTAGTGAGTG 386
QY 41 LeuAspGluLeuTyrrArgValThrArgLysLysThrGlnAsnLysLysGluAlaGluLys 60
Db : : : : :
387 CTGGATGAGCTCTACAGAGTACAGAGGAGTACCCCAAAACAAAGAGGAGGAGAG 446
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnGln 80
Db : : : : :
447 ATCATCAAGAACTCTATCAAGACATCATCAAGCTGGCCATTTTATAGAAATTAATCAG 506
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db : : : : :
507 TTTAAATCAAGATGAGCTAGCATGATGAGAGAAATTTAAGAAAGAAAGTTCAATCAGCTTGGCT 566
QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValIleuSerArg 120
Db : : : : :
567 ATACCGGTGCTCATGTTCCATCAGTGGATTAATACCTTGAACCGGAATGTGTTATCCAG 626
QY 121 LeuLeuAsnGluCysArgGluMetLysGlnIleIleGlnArgHisLeuThrAlaLys 140
Db : : : : :
627 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTGCAG 686
QY 141 SerHisGluValArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
Db : : : : :
687 TCCATGAGACGGGTAAATATGTTGATGATTTTTCAGATTGTGATTTTGGCTGCC 746
QY 161 LeuTyrrAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
Db : : : : :
747 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTCAAAAACATATGTGATGATATCAAC 806
QY 181 LysMetLeuAspGluGluAsnIle 188
Db : : : : :
807 AAAATGTTGATGAGAGACATATA 830
RESULT 14
AC035144
LOCUS AC035144 112626 bp DNA linear PRI 27-APR-2001

DEFINITION Homo sapiens chromosome 5 clone CTD-2220M12, complete sequence.
AC035144
AC035144.3 GI:13811899
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 112626)
DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL
TITLE
REFERENCE
2 (bases 1 to 112626)
DOE Joint Genome Institute.
JOURNAL
TITLE
REFERENCE
3 (bases 1 to 112626)
DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL
TITLE
AUTHORS
COMMENT
Direct Submission
Submitted (27-APR-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Apr 27, 2001 this sequence version replaced gi:7712094.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.2.
STS Content:
SHGC-7119 GI4211.
Location/Qualifiers
1. 112626
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2220M12"

ORIGIN

Alignment Scores:
Pred. No.: 1.4e-73 Length: 112626
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x AC035144 (1-112626)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
Db : : : : :
65048 GTGGCCACAGATGCTTTTAATCCAAAACCTGGCCCTTACGACACAAAAGAAATCTTG 65107
QY 21 GlyLysMetValSerLysSerIleAlaThrThreLeuIleAspAspThrSerSerGluVal 40
Db : : : : :
65108 GGTAAATGGTGTCCAAATCCATCGCCACACCTTAAATAGACACACAAAGTAGTGAGTG 65167
QY 41 LeuAspGluLeuTyrrArgValThrArgLysLysThrGlnAsnLysLysGluAlaGluLys 60
Db : : : : :
65168 CTGGATGAGCTCTACAGAGTACAGAGGAGTACCCCAAAACAAAGAGGAGGAGAG 65227
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnGln 80
Db : : : : :
65228 ATCATCAAGAACTCTATCAAGACATCAAGCTGCGCATTTTATAGAAATTAATCAG 65287
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db : : : : :
65288 TTTAATCAAGATGAGCTAGCATTTGAGAAATTTAAGAAAGTTTATCAGCTTGGCT 65347
QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValIleuSerArg 120
Db : : : : :
65348 ATGACCGGTGCTCATGTTCCATCAGTGGATTAATCCTTTGACCGGAATGTGTTATCAGG 65407

QY 121 LeuLeuAsnGluCyArGluMetLeuHISGlnIleIleGlnArGHisLeuThrAlaLys 140
|||||
Db 65408 CTGTTAAATGATGACGAGAGATGCTGCACCAATATTCAGGCGCACTCCTCAGG 65467
QY 141 SerHisGlyArGValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
65468 TCACATGACGAGGGTTAATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 65527
QY 161 LeuTyraAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
65528 TTGTAAATATCTTTTGGGAATTTTAAACCCACTTCAAAACTATGATGATCAAC 65587
QY 181 LysMetLeuAspGluGluAsnIle 188
65588 AAAATTTGATGAGAGAACATTA 65611
Db
RESULT 15
AC026795/c 156277 bp DNA linear PRI 21-SEP-2001
LOCUS Homo sapiens chromosome 5 clone CTD-2287L22, complete sequence.
DEFINITION AC026795
AC026795.6 GI:15718543
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL 2 (bases 1 to 156277)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL 2 (bases 1 to 156277)
AUTHORS Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
JOURNAL 3 (bases 1 to 156277)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
JOURNAL On Sep 21, 2001 this sequence version replaced gi:15383783.
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.4.
STS Content:
SHGC-7119 G14211.
FEATURES
source Location/Qualifiers
1..156277
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2287L22"

ORIGIN

Alignment Scores:
Pred. No.: 1.96e-73 Length: 156277
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x AC026795 (1-156277)
QY 1 MetAlaThrAspValPheAsnSerIysPheLeuAlaValGlnAlaGlnLysIleLeu 20
:::
Db 128234 GTGGCCACAGATGCTTTAATATCCAAAACCTGGCCGTTCAGGCACAAAAGAAATCTTG 128175
QY 21 GlyMetValSerIysSerIleAlaThrThrLeuIleAspThrSerSerGluVal 40
128174 GGTAAATATGCTCCAAATCCATGCCCAACCTTAATAGACACAAAGTAGTAGGGT 128115
QY 41 LeuAspGluLeuTyraGValThrArgLysTyThrGlnAsnLysGlnAlaGluLys 60
128114 CTGATGAGCTCTTACAGAGTGACAGGAGTACACCCAAAACAAAGAGGACAGAAAG 128055
QY 61 LysIleLysAsnLeuIleIleThrValIleLysLeuAlaIleLeuTyraGsnAsnGln 80
128054 ATCATCAAGAACTCATCAAGCACTCATCAAGCTGGCCATCTTTATAGAAATATCAAG 127995
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db 127994 TTTAATCAAGATGAGCTAGCTATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT 127935
QY 101 MetThrValAlaSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
127934 ATGACCGGTGTCAGTTCCATCAGGTGATTTATACCTTGAACCGGAATGTGTTATCCAGG 127875
QY 121 LeuLeuAsnGluCyArGluMetLeuHISGlnIleIleGlnArGHisLeuThrAlaLys 140
Db 127874 CTGTTAAATGAATGACGAGAGATGCTGCACCAATATTCAGGCGCACTCCTCAGG 127815
QY 141 SerHisGlyArGValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
Db 127814 TCACATGACGAGGGTTAATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 127755
QY 161 LeuTyraAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
Db 127754 TTGTAAATATCTTTTGGGAATTTTAAACCCACTTCAAAACTATGATGATCAAC 127695
QY 181 LysMetLeuAspGluGluAsnIle 188
Db 127694 AAAATTTGATGAGAGAACATTA 127671
Db
RESULT 16
AC138612/c 158057 bp DNA linear PRI 30-MAY-2003
LOCUS Homo sapiens BAC clone CTD-223E22 from 2, complete sequence.
DEFINITION AC138612
AC138612.1 GI:27597054
ACCESSION
VERSION AC138612.1 GI:27597054
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Toward a complete human genome sequence
JOURNAL Submitted (11-JAN-2003) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
JOURNAL 4 (bases 1 to 158057)
AUTHORS Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-2003) Department of Genetics, Washington

COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

----- Genome Center
Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@watson.wustl.edu

----- Summary Statistics

Center project name: H_MS2293E22

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTD-2293E22 is from a release of the human BAC library CTD. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
VECTOR: pBelOBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

```

source
    1..158057
       /organism="Homo sapiens"
       /mol_type="genomic DNA"
       /db_xref="taxon:9606"
       /chromosome="2"
       /map="2"
       /clone="CTD-2293E22"
       /clone_1b="CTD"
repeat_region
    4..83
       /rpt_family="MIR"
repeat_region
    300..719
       /rpt_family="ERV1"
repeat_region
    1321..1459
       /rpt_family="L1"
repeat_region
    1518..1547
       /rpt_family="AT_rich"
repeat_region
    1964..1996
       /rpt_family="(TA)n"
repeat_region
    2532..2575
       /rpt_family="A-rich"
repeat_region
    2639..2841
       /rpt_family="L2"
repeat_region
    2859..2886
       /rpt_family="AT_rich"
repeat_region
    3179..3201
       /rpt_family="AT_rich"
repeat_region
    4351..4541
       /rpt_family="MIR"
repeat_region
    5128..5766
       /rpt_family="L1"
repeat_region
    5767..6328
       /rpt_family="L1"
repeat_region
    6329..7499
       /rpt_family="L1"
repeat_region
    7503..7872
       /rpt_family="L1"
repeat_region
    7920..8106
       /rpt_family="MALR"
repeat_region
    8621..8849
       /rpt_family="ERV1"
repeat_region
    9113..9415
       /rpt_family="Alu"
repeat_region
    12251..12442
       /rpt_family="L1"
repeat_region
    14323..14361
       /rpt_family="AT_rich"
repeat_region
    14630..14899
       /rpt_family="Alu"
repeat_region
    15273..15541
       /rpt_family="L1"
repeat_region
    15534..17744
       /rpt_family="L1"
repeat_region
    18238..18436
       /rpt_family="MER2_type"
repeat_region
    18725..18973
       /rpt_family="MER2_type"
repeat_region
    19548..19664
       /rpt_family="L2"
repeat_region
    19666..19723
       /rpt_family="L2"
repeat_region
    20144..20192
       /rpt_family="TAAAG"
repeat_region
    20415..20468
       /rpt_family="polypurine"
repeat_region
    20476..20771
       /rpt_family="AT_rich"
repeat_region
    21591..22250
       /rpt_family="Alu"
repeat_region
    22475..22483
       /rpt_family="L1"
repeat_region
    24487..24515
       /rpt_family="AT_rich"
repeat_region
    25137..25196
       /rpt_family="CA"
repeat_region
    26530..26604
       /rpt_family="L2"
repeat_region
    27433..28325
       /rpt_family="L2"
repeat_region
    28361..28469
       /rpt_family="MIR"
repeat_region
    28603..28653
       /rpt_family="(CATATA)n"
repeat_region
    29092..29113
       /rpt_family="AT_rich"
repeat_region
    29379..29680
       /rpt_family="Alu"
repeat_region
    29871..30008
       /rpt_family="MER1_type"
repeat_region
    30569..30609
       /note="Sequence derived from PCR product of project DNA."
repeat_region
    31531..31588
       /rpt_family="GA-rich"
repeat_region
    31939..32107
       /rpt_family="ERV1"
repeat_region
    32426..32605
       /rpt_family="Alu"
repeat_region
    32814..32878
       /rpt_family="MIR"
repeat_region
    33458..33570
       /rpt_family="L2"
repeat_region
    33887..33984
       /rpt_family="(GAAA)n"
repeat_region
    35523..35595
       /rpt_family="AT_rich"
repeat_region
    36412..36602

```

```

repeat_region      /rpt_family="MIR"
37311..37413
repeat_region      /rpt_family="L1"
37414..37724
repeat_region      /rpt_family="Alu"
37725..37879
repeat_region      /rpt_family="L1"
38402..38429
repeat_region      /rpt_family="AT_rich"
38459..38494
repeat_region      /rpt_family="AT_rich"
39513..39548
repeat_region      /rpt_family="cRNA-Glu-GAG"
39654..39929
repeat_region      /rpt_family="L1"
39956..40148
repeat_region      /rpt_family="MIR"
41023..41048
repeat_region      /rpt_family="(A)n"
41243..41431
repeat_region      /rpt_family="L1"
41475..41561
repeat_region      /rpt_family="L1"
41789..41811
repeat_region      /rpt_family="AT_rich"

Alignment Scores:
Pred. No.:      1 99e-73      Length:      158057
Score:          950.00      Matches:      186
Percent Similarity: 99.47%      Conservative: 1
Best Local Similarity: 98.94%      Mismatches: 1
Query Match:    98.75%      Indels:      0
DB:             9      Gaps:      0

US-10-627-571-2 (1-188) x AC138612 (1-158057)

QY      1 MetAlaThrAspValPheAsnSerIlyAsnLeuAlaValGlnAglInySylLeu 20
Db      135468 GGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTTGAGCACAAGAAAGATCTTG 135409
QY      21 GILySMeCValSerIySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db      135408 GGTAAATGATGTCGCAATCCATCGCACCACTTAATAGACACAAAGTGTGAGTGG 135349
QY      41 LeuAspGluLeuThrArgValThrArgGluThrThrGlnAsnIlySglValGluVal 60
Db      135348 CTGGATGAGCTCTCAAGAGTACAGAGTACACCCAAACAAAGAGGCGACGAGAG 135289
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuThrArgAsnAngln 80
Db      135288 ATCATCAAGAACCTCATCAAGACAGTCACTCAAGCTGGCATTTCTTATAGAAATATCAG 135229
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIlySylValHisGlnLeuAla 100
Db      135228 TTTAATCAAGATGAGCTAGATTTGATGAGAAATTTAAGAAAGTTCATCAGTTGCT 135169
QY      101 MetThrValValSerPheHisGlnValAspIlyThrPheAspArgAsnValLeuSerArg 120
Db      135168 ATGACCGTGGTCAGTTTCCATCAAGTGGATTAACCTTTGACCGGAATGTGTATCCAG 135109
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db      135108 CTGTTAAATGATGACGAGAGAGTGCACCAATATTCATCAGCGCACCTCATGCGCAAG 135049
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
Db      135048 TCACATGAGACGGGTATATATGTTGATCATTTTTCAGATTTGTAATTTTGGCTGCC 134989
QY      161 LeuThrAspProPheGlyAsnPheLysProHisLeuGlnIlyLeuGlyAspArgIlyIleAsn 180
Db      134988 TTGTATTAATTCCTTTTGGAAATTTTAAACCCACTTACAAAACTAATGTGATGATCAAC 134929
QY      181 LysMetLeuAspGluGluAsnIle 188

```

```

Db      134928 AAAATGTCGATGAGAGACACTA 134905
RESULT 17
BD135356
LOCUS      BD135356      1943 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION 110 human secretory proteins.
ACCESSION  BD135356
VERSION     BD135356.1 GI:23230301
KEYWORDS   JP 2002508167-A/107.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  Moore,P.A., Ruben,S.M., Carter,K.C., Shi,Y., Rosen,C.A.,
            Soppet,D.R., Caou,H., Wei,Y.F., Florence,K., Duan,R.D.,
            Florence,C., Greene,J.M., Feng,P., Ferrite,A.M., Yu,G.L., Janat,F.
            and Ni,J.
            110 human secretory proteins
            Patent: JP 2002508167-A 107 19-MAR-2002;
TITLE      JOURNAL
COMMENT    HUMAN GENOME SCIENCES INC
            OS Homo sapiens (human)
            PN JP 2002508167-A/107
            PD 19-MAR-2002
            PF 17-DEC-1998 JP 2000539040
            PR 18-DEC-1997 US 60/068 057,18-DEC-1997 US 60/068 007 PR
            18-DEC-1997 US 60/068 008,18-DEC-1997 US 60/068 054 PR
            18-DEC-1997 US 60/068 064,18-DEC-1997 US 60/068 053 PR
            19-DEC-1997 US 60/068 169,19-DEC-1997 US 60/068 368 PR
            19-DEC-1997 US 60/068 367,19-DEC-1997 US 60/068 369 PR
            19-DEC-1997 US 60/068 365
            PI PAUL A MOORE,STEVEN M RUBEN,KENNETH C CARTER,YANGU SHI,CRAIG
            PI A ROSEN,
            PI DANIEL R SOPPET,HARA CAOU,YING FEI WEI,KIMBERLY FLORENCE, PI
            ROSANNE D DUAN,
            PI CHARLES FLORENCE,JOHN M GREENE,PING FENG,ANN M FERRIE,GUO PI
            LIANG YU
            PI FORD JANAT,JIAN NI
            PC C12N15/09,A61K38/00,A61K48/00,A61P9/00,A61P9/10,A61P15/00, PC
            A61P25/00,
            PC A61P25/02,A61P25/14,A61P25/16,A61P25/18,A61P25/22,A61P25/24,
            PC A61P25/28,
            PC A61P29/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
            10,
            PC C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53,C12N15/00, PC
            A61K37/02,
            PC C12N5/00
            CC n equals a,t,g,or c
            CC n equals a,t,g,or c
            CC n equals a,t,g,or c
            CC n equals a,t,g,or c
            CC n equals a,t,g,or c
            FH Key
            FT source
            FT location/Qualifiers
            1..1943
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      3 21e-75      Length:      1943
Score:          948.00      Matches:      186
Percent Similarity: 98.94%      Conservative: 0
Best Local Similarity: 98.94%      Mismatches: 2
Query Match:    98.54%      Indels:      0
DB:             6      Gaps:      0

US-10-627-571-2 (1-188) x BD135356 (1-1943)

```

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAglInlySylleu 20
DB      120 ATGGCCACAGATATCTTTAATTCAAAACCTGCGCTTANCCACAAAAGAAATCTTG 179
QY      21 GlylMetValSerIysSerIleAlaThrThleuIleAspAspThrSerSerGluVal 40
DB      180 GGTAAATGCTGTCCAATTCATCGCCACACCTTATAGACACACAGATGAGGTG 239
QY      41 LeuAspGluLeuTyArgValThrArgIuTyThrGlnAsnIysGluAgluIys 60
DB      240 CTGGATGAGCTTACAGAGTACAGAGGACGAGGAGTACACCCAAAACAGAGGAGGAG 299
QY      61 LysIleIysAsnLeuIleIysThrValIleIysLeuAlaIleIysTyArgAsnGln 80
DB      300 ATCATAGAGAACTCATCAAGACAGATCAACCTGCGCTTCTTATAGAAATTAATCAG 359
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIysIysValIhIsgInleuAla 100
DB      360 TTTAAATCAAGATAGCTAGCATTTGATGAGAAATTTTAAAGAAAGTTTCATCAGCTTGT 419
QY      101 MetThrValIserPheIhIsgInValAspTyThrPheAspArgAsnValLeuSerArg 120
DB      420 ATGACCGTGGTCACTTCCATCAGGTGAGTATACCTTTCACCGAATGTATTATCCAGG 479
QY      121 LeuLeuAsnGluCysArgIuMetLeuIhIsgInIleIleGlnArgIhIleuThrAlaIys 140
DB      480 CTGTTAAATGAATGACAGAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTCCAG 539
QY      141 SerhIglYArgValAsnAsnValPheAspIhPheSerAspGluPheLeuAlaIa 160
DB      540 TCCATAGAGAGGGATTAATATGTTGATCATTTTCAGATTGTGAATTTTGCGCTGCC 599
QY      161 LeuTyArgAspProPheGlyAsnPhelysProhIleuGlnIysLeuGlyAspGlyIleAsn 180
DB      600 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTCAAAAATATATGTATGTATATC 659
QY      181 LysMetLeuAspGluGluAsnIle 188
DB      660 AAAATGTTGATGAGAGAACATTA 683

```

RESULT 18
 AK097284
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ39965 fis, clone SPLEN2027157, highly similar to Homo sapiens MDC-3.13 isoform 2 mRNA.
 ACCESSION AK097284
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, Y., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirakawa, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Hosoda, M., Hootu, T., Kaneko, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shichata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,

Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiro, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL
 PUBMED
 14702039

REFERENCE
 AUTHORS
 Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S., Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hootu, T., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kametani, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 Location/Qualifiers
 FEATURES
 source
 1..2502
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="SPLEN2027157"
 /tissue_type="spleen"
 /clone_lib="SPLEN2"
 /note="cloning vector: pME18FL3"

ORIGIN
 Alignment Scores:
 Pred. No.: 9,486-75 Length: 2502
 Score: 944.00 Matches: 185
 Percent Similarity: 99.478 Conservative: 2
 Best Local Similarity: 98.408 Mismatches: 1
 Query Match: 98.134 Indels: 0
 DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x AK097284 (1-2502)

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAglInlySylleu 20
DB      120 ATGGCCACAGATATCTTTAATTCAAAACCTGCGCTTANCCACAAAAGAAATCTTG 179
QY      21 GlylMetValSerIysSerIleAlaThrThleuIleAspAspThrSerSerGluVal 40
DB      180 GGTAAATGCTGTCCAATTCATCGCCACACCTTATAGACACACAGATGAGGTG 239
QY      41 LeuAspGluLeuTyArgValThrArgIuTyThrGlnAsnIysGluAgluIys 60
DB      240 CTGGATGAGCTTACAGAGTACAGAGGACGAGGAGTACACCCAAAACAGAGGAGGAG 299
QY      61 LysIleIysAsnLeuIleIysThrValIleIysLeuAlaIleIysTyArgAsnGln 80

```

```

Db      899 ATCATCAAGAACTCATCAAGACAGCATCAAGCTGCGCATTTTATAGAAATATACG 958
Qy      81 Pheasnglaspgluleu1aleuwtgltulyspheylalye1ahi1sgluleu1a 100
Db      959 TTTAATCAAGATGAGCTACATTTGAAGAAATTTAAGAGAAATTCATCAGCTTGCT 1018
Qy      101 MettThValSerpHeh1eg1nValaspYrThPhasapArganValleuSerArg 120
Db      1019 AAGACCGTGTGAGTTTCCATCAATGAGTGAGATTAATCCTTTGACCGGATGTGATCCAG 1078
Qy      121 LeuleasnglucYsarYsgluleuH1sg1nle1le1gnarYh1sDeutHrAla1ys 140
Db      1079 CTTTAAATGAATGACGAGAGAGATGCTGTACAAATCATTCAGGCCACTCCTGCGAAG 1138
Qy      141 SerH1sglyArYValasnAnValPheasPh1sPheSerapYsg1nPhelueu1a1a 160
Db      1139 TCACATGAGCGGTTTAAATGCTTTGATCATTTTTCAGATTTGAAATTTTGCGCTGC 1198
Qy      161 LeuYrAspProPhelYasPhelysProH1sleuG1lyleuYsaPpYl1leu1 180
Db      1199 TTGTAAATCCTTTTGGAAATTTTAAACCCCACTTCAAAAATGATGATGATCAAC 1258
Qy      181 LysMeleuasPgluGluAsn1le 188
Db      1259 AAAATGTTGATGAAGAGAAACATA 1282

RESULT 19
LOCUS   BC009090              1728 bp      mRNA      linear      ROD 30-JUN-2004
DEFINITION Mus musculus tumor necrosis factor, alpha-induced protein 8, mRNA
ACCESSION BC009090
VERSION   BC009090.1 GI:14318585
KEYWORDS  MGC.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1728)
          Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
          Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
          Altschul,S.F., Zeeberg,B., Buetow,K.H., Scheefer,C.F., Bhat,N.K.,
          Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
          Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
          Stapleton,M., Soares,M.B., Bonaldo,M.F., Caeavanc,T.L.,
          Scheetz,T.E., Brownstein,M.J., Useth,T.B., Toshiyuki,S.,
          Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
          Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
          McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
          Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
          Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
          Fahy,U., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
          Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
          Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
          Dickson,M.C., Rodriguez,A.C., Grumwood,J., Schmitt,J., Myers,R.M.,
          Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallue,D.B.,
          Schermer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL PUMED 12477932
REFERENCE 2 (bases 1 to 1728)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
          Email: cgaabs-r@mail.nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.

```

```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
          Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
          Kowals, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nananvati,
          A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAK Place: 17 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19527361.
Location/Qualifiers
1..1728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:11714 IMAGE:3965693"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
1..1728
/gene="tnfaip8"
/notes="synonyms: Nded, Gq2-1, Ssc-2"
/db_xref="locusid:106869"
/db_xref="MGI:2147191"
70..666
/gene="tnfaip8"
/codon_start=1
/product="TNF-induced protein"
/protein_id="AAH09090.1"
/db_xref="GI:14318586"
/db_xref="MGI:2147191"
/translacion="M1S2AEPREAVATDFVNSKNIYAQOKKILGKMSKSIATTLID
DTSEVLELVKTVSTYONKKEARVKNIKITVYKLVLRNNQPNODELALMEKF
KKYKQLMALVTSFHVQVETPTDRNVLSTLNLCERELHEIIRHILTAHSRGNVNVFP
HFSDDFLAALYNPRKFRPHQKLCIDGINKMLDEBNT"

ORIGIN
Alignment Scores:
Pred. No.: 8,966-73 Length: 1728
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 10 Gaps: 0

US-10-627-571-2 (1-188) x BC009090 (1-1728)
Qy      1 MetAlaThrAspValPheasnserylsanleu1a1a1GlnAlaGlnlyls1leu 20
Db      100 GTGGCTACAGATGCTTCATTCAAAAACCTGGCGCTTCAGGCACAAAGAAATCCCG 159
Qy      21 GlylYMeleValSerlySer1leAlaThrThrleu1leaspAspThSerSerGluVal 40
Db      160 GGCAAAATGGATCCAAATCCATCCGACACAGCTGATCGACACACAGCAGAGGTG 219
Qy      41 LeuAspGluLeuYrArgValThraArgluYrThrGlnAsnlylysg1nlaGluYs 60
Db      220 CTAGATGAGCTTACAGGCTGACCAAGAGTACACCCAGAACAGAGAGCGGAGAGG 279
Qy      61 Lys1lelyAsnleu1lelysrThrVal1lelyleu1a1le1eulYrArgAsnAng1n 80
Db      280 GTCATCAAGAACTCATCAAGACGCTCATCAAGCTGCGCTCTCCACAGAGAAATACAG 339

```



```

/rpt_family="B4A"
repeat_region 3506.3534
/rpt_family="(CA)n"
complement(4228.4578)
repeat_region /rpt_family="WTE"
4848.5049
repeat_region /rpt_family="B3"
5526.5572
/rpt_family="AT_rich"
6521.6815
/rpt_family="Lx9"
6890.6921
/rpt_family="AT_rich"
8693.8735
/rpt_family="(TTCC)n"
8699.8700
/notes="probably CC"
repeat_region 9208.9234
/rpt_family="(TTTG)n"
complement(9236.9335)
/rpt_family="PB1D7"
complement(9603.9751)
repeat_region /rpt_family="B1F"
10571.10599
/rpt_family="AT_rich"
10637.10660
/rpt_family="AT_rich"
10951.11146
/rpt_family="B2_Mm1"
12671.12730
/rpt_family="(TG)n"
13141.13214
/rpt_family="ID3"
13228.13256
/rpt_family="PB1D10"
13257.13323
/rpt_family="ID_B1"
13331.13358
/rpt_family="AT_rich"
13483.13667
/rpt_family="B3"
13769.14378
/rpt_family="L1M4"
14519.14619
/rpt_family="L1_MM"
complement(14772.15069)
/rpt_family="L1_MM"
15389.15420
/rpt_family="AT_rich"
17608.17670
/rpt_family="(TATATG)n"
17773.17801
/rpt_family="AT_rich"
17825.17850
/rpt_family="AT_rich"
18110.18163
/rpt_family="AT_rich"
20210.20731
/rpt_family="L1_MM"
20759.20876
/rpt_family="RSINB1"
complement(21592.21799)
/rpt_family="B1_MM"
21818.21837
/rpt_family="(A)n"
21916.21949
/rpt_family="AT_rich"
complement(22545.22831)
/rpt_family="Lx7"
22892.23011
/rpt_family="B1_MM"
22949.22974
/notes="single clone coverage"

```

```

repeat_region complement(23095.23294)
/rpt_family="Lx7"
repeat_region complement(23912.24049)
/rpt_family="B1_MM"
repeat_region complement(24017.24236)
/rpt_family="MTC"
repeat_region complement(24238.24355)

Alignment Scores:
Pred. No.: 1,11e-70 Length: 184327
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 10 Gaps: 0

US-10-627-571-2 (1-188) x AC120859 (1-184327)

QY 1 MetAlaTrnAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeu 20
...
DB 144853 GTGGCTACAGATGCTTCATTCCAAAAACCTGGCGTTCCAGGCACAAAAGAGATCCTG 144912

QY 21 GlyLysMetValSerLysSerLLeaLathrThrLeuLLeaSPRThSerSerGluVal 40
...
DB 144913 GGCAAAATGCTATCCAAATCCATCCGCCACCGCTGATCGACACACGACGCGAGGTG 144972

QY 41 LeuAspGluLeuTyraArgValThrArgGluTyrrThrgLnaenLysGluAlaGluLys 60
...
DB 144973 CTAGATGAGCTGTACAGGCTGACCAAGATGACCCAGAACACAAAGAGCGGAGAGG 145032

QY 61 LysIleLysAsnLeuLLeuLysThrValLLeuLysLeuAlaLLeuTyraArgAsnGln 80
...
DB 145033 GTCATCAAGAACCTCATCAAGACGCTCATCAAGCTGGCGTCTTCACACGAAACATCG 145092

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
...
DB 145093 TTCAAATCAAGACGAGCTGGCGCTCATGAGAAATTCAGAAAGAGTGCCACGCTTGGC 145152

QY 101 MetTrnValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
...
DB 145153 ATGACGGTGTGATGCTTCACACCGATGAGTACACCTTCGACCGCAATGTCTGTCCAGG 145212

QY 121 LeuLeuAsnGluCyArgGluMetLeuHisGlnLLeuLLeuArgHisLeuThrAlaLys 140
...
DB 145213 CTGGCTGAACGAGTGGCGAGAGCTCCTACACGAGATCATTCAGCGCACCTTACCGCCAAG 145272

QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
...
DB 145272 TCTCACGCGACGGGTTAATATATGTCTTTCAGCATTTTTCAGATTGATTTTGGCTGCGC 145332

QY 161 LeuTyraenProPheGluAsnPhelyProHisLeuGlnLysLeuCyAspGluLLeuAsn 180
...
DB 145333 TTGTCAATCCCTTTGGAAAGTTTAACCTCATTCAGAAACCTTTCGACGCGCATCAAC 145392

QY 181 LysMetLeuAspGluGluAsnLLeu 188
...
DB 145393 AAAATGTTAGATGAAGAGAACATA 145416

RESULT 21
AC095257 237561 bp DNA linear HTG 09-NOV-2002
LOCUS Rattus norvegicus clone CH230-10A12, WORKING DRAFT SEQUENCE.
AC095257
AC095257.5 GI:24818022
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 237561)
REFERENCE
1 Muzny,D.Marie., Metzker,M.Iee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Angilano,D.,

```

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Day-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denon, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gibrageorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howell, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, S., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, D., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Louisse, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Ming, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwachukwu, O., Okunolu, G., Olarnunpasoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Piopier, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Weinstock, G., and Gibbs, R. A.

Unpublished
Direct Submission
2 (bases 1 to 237561)
Worley, R. C.
Direct Submission
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 237561)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23264563.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GDFL
Center clone name: CH230-10A12

Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 229749 bases at least Q40
Consensus quality: 231209 bases at least Q30
Consensus quality: 232003 bases at least Q20
Estimated insert size: 232575; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 237561: contig of 237561 bp in length.
Location/Qualifiers
1. 237561
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-10A12"
1. 1156
/note="wgs end-extension
clone_end:Sp6
/note="wgs end-extension
clone_end:Sp6
/note="clone boundary
clone_end:Sp6
site:BCORI
end_sequence:BH03075"
76320..79576
/note="wgs contig"
complement(23242..233289)
/note="clone boundary
clone_end:T7
site:BCORI
end_sequence:BH03071"
234103..235440
/note="wgs end-extension
clone_end:T7"
235793..237561
/note="wgs end-extension
clone_end:T7"

ORIGIN
Alignment Scores:
Pred. No.: 2,17e-70 Length: 237561
Score: 918.00 Matches: 177
Percent Similarity: 98.40% Conservative: 8
Best Local Similarity: 94.15% Mismatches: 3
Query Match: 95.43% Indels: 0
DB: 2 Gaps: 0

US-10-627-571-2 (1-188) x AC095257 (1-237561)

QY 1 MetcAlaThrAspValPheAsnSerLysAsnLeuAlaGlnAlaGlnLysLysIleLeu 20
:::|||||
Db 95045 GAGCGTACAGAGTCTTTAATTCACAAAACCTGGCGCTTCAGGACAAAAGAGATCTCG 95104
QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40

```

Db      95105 GGTAAATGATCCAAATCCATCGCCACCACTTGATGATGACGACGAGCGAGTGC 95164
Qy      41 LeuAaPGLuLeuTyraArgValThrArgGluTyrrThGlnAsnLysLysGlnAgluLys 60
Db      95165 CTGGAGAGACTGTACAGGGGTGACCAAGAGTACACCAAGAAAGAGGAGCGAGGAAG 95224
Qy      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnAngin 80
Db      95225 GTCATCAAGAACTCTCATCAAGAGGTCACTCAAGCTGGCGTTTCCACGAGAACTACG 95284
Qy      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db      95285 TTCACACCAAGACGAGCTGCTTGATGAGAGAGTTCAAGAAAGGTGACCGAGCTGGCC 95344
Qy      101 MetThrValValSerPheHisGlnValAspTyrrThPheAspArgAsnValLeuSerArg 120
Db      95345 ATGACCGGTGTAGCTTTCACCAAGGTAGTACACCTTTGACCGCAACGTGCTGCAGG 95404
Qy      121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db      95405 CTGCTGAACGAGTGGCCGAGAGCTCTGACGACATCATTCAGGCCACCTTACCGCCAG 95464
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
Db      95465 TCTCAGCGAGCGGGTAAATACGCTTGTGATCATTTTTCAGATTGTGATTTTGGCTGCC 95524
Qy      161 LeuTyraAsnProPheGluAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
Db      95525 TTGTAAATATCCCTTGGAAAATTTAAACCTTACAGAACTTGTGACGGCATCAAC 95584
Qy      181 LysMetLeuAspGluGluAsnIle 188
Db      95585 AAAATGTTGGATGAAGAAACATA 95608

RESULT 22
BD149395      816 bp      DNA      linear      PAT 17-JAN-2003
LOCUS      BD149395      Primer for synthesizing full-length cDNA and use thereof.
DEFINITION      BD149395
ACCESSION      BD149395.1 GI:27855153
VERSION      JP 2002191363-A/4238.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE      Primer for synthesizing full-length cDNA and use thereof
JOURNAL      Patent: JP 2002191363-A 4238 09-JUL-2002;
COMMENT      HELIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/4238
PD      09-JUL-2002
PF      28-JUL-2000 JP 2000280990
PI      TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAZUO
PI      SAITO.
PI      JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI      KEIICHI NAGAI, TETSUO OTSUKI
PC      C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
PC      10, C12P2/02, C12Q1/68//C12P2/08, G06F17/30, C12N15/00, C12N5/00 CC
PC      Primer for synthesizing full-length cDNA and use thereof FH Key
FT      source      1. 816
FT      Location/Qualifiers
FT      1. 816
FT      /organism='Homo sapiens (human)'.
FEATURES
source      1. 816
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

```

```

ORIGIN
Alignment Scores:
Pred. No.:      9,416-73      Length:      816
Score:      916.00      Matches:      184
Percent Similarity:      97.89%      Conservative:      2
Best Local Similarity:      96.84%      Mismatches:      2
Query Match:      95.22%      Indels:      2
DB:      6      Gaps:      0

US-10-627-571-2 (1-188) x BD149395 (1-816)
Qy      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
Db      101 GTGGCCACAGAGTCTTTAATTCGAAAAACCTGGCGCTTCCAGGCACAAAGAGATCTTG 160
Qy      21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db      161 GGTAAATGCTGTCCAAATCCATCCGCCACCACTTAATAGACGACACAAAGTAGTAGGTG 220
Qy      41 LeuAaPGLuLeuTyraArgValThrArgGluTyrrThGlnAsnLysLysGlnAgluLys 60
Db      221 CTGGATGAGCTTACAGAGTGAACAGGAGTACACCAAAACAAAGAGAGGACAGGAAG 280
Qy      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnAngin 80
Db      281 ATCATCAAGAACTCTCATCAAGACAGTCAAGCTGGCCATTCTTTATAGAAATATCAG 340
Qy      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db      341 TTTAATCAAGAGAGACTGATGATGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGCT 400
Qy      101 MetThrValValSerPheHisGlnValAspTyrrThPheAspArgAsnValLeuSerArg 120
Db      401 ATGACCGGTGTAGTTCATCAGAGTGAATTAATCTTTGACCGGAAATGTGATTCAGG 460
Qy      121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db      461 CTGTTAATATGATGACGAGAGATGCTGCAACCAATCATTCAGCGCCACCTCAGCAAG 520
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
Db      521 TCACATGAGCGGGTAAATATGATGTTGATCATCTTTTTCAGATTGTGAAATTTGGCTGCC 580
Qy      161 LeuTyraAsnProPheGluAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
Db      581 TTGTAAATATCCCTTGGGAAATTTAAACCCCACTTACAAAACATAAGTAGATGATCAAC 640
Qy      181 LysMet--LeuAaPGLuGluAsnIle 188
Db      641 AAAATGTTGGATGAAGAAACATA 666

RESULT 23
AX869333      816 bp      DNA      linear      PAT 17-DEC-2003
LOCUS      AX869333      Sequence 4238 from Patent EP1074617.
DEFINITION      AX869333
ACCESSION      AX869333
VERSION      AX869333.1 GI:40024196
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE      Primers for synthesizing full-length cDNA and their use
JOURNAL      Patent: EP 1074617-A 4238 07-FEB-2001;
COMMENT      Research Association for Biotechnology (JRP)
FT      source      1. 816
FT      Location/Qualifiers
FEATURES
source      1. 816
/organism='Homo sapiens'
/mol_type='unassigned DNA'

```


ORIGIN /db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 9,41e-73 Length: 816
Score: 916.00 Matches: 184
Percent Similarity: 97.89% Conservative: 2
Best Local Similarity: 96.84% Mismatches: 2
Query Match: 95.22% Indels: 2
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x AX869333 (1-816)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAglInuLysIleLeu 20
DB 101 GTGGCCACAGATGCTTAAATTCACAAAACCTGGCCCTTACAGCACAAGAGATCTTG 160
QY 21 GlyMetValSerLysSerIleAlaThrThreuleAspThrSerSerGluVal 40
DB 161 GGTAAATGGTGTCCAAATCCATGCGACCACTTAATAGCGACAGAGAGAGAG 220
QY 41 LeuAspGluLeuTyraGValThraGluTyrrThrGlnAsnLysGluAgluLys 60
DB 221 CTGGATGAGCTCTACAGAGTACCGAGGAGTACCCAAAACAAGAGAGAGAG 280
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraGAsnGln 80
DB 281 ATCATCAAGAACCTCATCAAGACAGTCAAGCTGCGCATTTCTTAAGAAATATCAG 340
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAla 100
DB 341 TTTAATCAAGATGAGTCAAGATGATGAGAAATTAAGAAAGTTCATCAGCTTGT 400
QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
DB 401 ATGACCGTGTGATGCTTCCATCAGTGCATTAATACCTTACCGAGATGCTTATCCAG 460
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleGlnArgHisLeuThrAlaLys 140
DB 461 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTAAGCGCCACCTCAGTGCAG 520
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerArgGluPheLeuAla 160
DB 521 TCACATGAGACGGGTAAATATGTTGATCATTTTCAATGTTGAATTTTGGCTGCC 580
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCyAspGlyLeuAsn 180
DB 581 TTTGATTAATCTCTTTGGGAATTTTAAACCCCATTCACAAACTAAGTGAATGATATCAC 640
QY 181 LysMet-LeuAspGluGluAsnIle 188
DB 641 AAAATGTTTGGATGAAGAACACTA 666

RESULT 24
AJ720906 2009 bp mRNA linear VRT 30-SEP-2004
LOCUS Gallus gallus mRNA for hypothetical protein, clone 29h8.
DEFINITION AJ720906
ACCESSION AJ720906.1 GI:53136471
VERSION
KEYWORDS ORF1.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE
AUTHORS Caldwell,R.B., Kierzek,A.M., Arakawa,H., Bezubov,Y., Zaim,J.,
Friedler,P., Kuter,S., Blagoderaki,A., Kostovska,D., Koter,M.,
Plachy,J., Carninci,P., Hayashizaki,Y. and Buerstedde,U.M.
TITLE Full-length cDNAs from bursal lymphocytes to facilitate gene
function analysis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2009)

AUTHORS Caldwell,R.B.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuberg, GERMANY
Location/Qualifiers

FEATURES
source 1..2009
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="29h8"
/cell_type="lymphocyte"
/tissue_type="bursa"
/clone_lib="rikeni"
/dev_stage="2 weeks old"
1..2009
/locus_tag="RCJMB04_29h8"
1..188
/locus_tag="RCJMB04_29h8"
189..755
/locus_tag="RCJMB04_29h8"
/note="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAG32565.1"
/db_xref="GI:53136472"
/translation="MATDVFNKSLAIQAKKILGRVSKSIATTLIDPTSDVDEL
VYRTKEYTONKKEAEKIKNLKIYKIALILVRNNOFNDETALMEKPKYHQIAK
VNSFHOVDTPRNPLSKLINDCRELIHQIRHLTAKSHGRVNVFDFRSDERFLAA
LYNPFEPYKLHLQKLCGVNRMIDSENT"
756..>2009
/locus_tag="RCJMB04_29h8"

ORIGIN
Alignment Scores:
Pred. No.: 3,88e-69 Length: 2009
Score: 880.00 Matches: 169
Percent Similarity: 96.28% Conservative: 12
Best Local Similarity: 89.89% Mismatches: 7
Query Match: 91.48% Indels: 0
DB: 5 Gaps: 0

US-10-627-571-2 (1-188) x AJ720906 (1-2009)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAglInuLysIleLeu 20
DB 189 ATGGCCACGAGATGCTTCAATTCACAAAAGCTTGGCCATTCAGGCCACGAAAGATCTT 248
QY 21 GlyLysMetValSerLysSerIleAlaThrThreuleAspThrSerSerGluVal 40
DB 249 GGGAAATGGTATCCAAAGTCAATGACATCACTTATGATGATGATATCCAGCACTGATGT 308
QY 41 LeuAspGluLeuTyraGValThraGluTyrrThrGlnAsnLysGluAgluLys 60
DB 309 TTGATGAGCTCTACAGAGTACGAGGAATACCAACCAAAACAAGAGAGAGAG 368
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraGAsnGln 80
DB 369 ATCATTAANAACCTCATTAATAATAGTCTCCAAATTTGGCAATTTCTTACAGAAACAATCA 428
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAla 100
DB 429 TTTAATCAAGATGAATTAATGACATGAGGAAGTTCACAGAAAGTTCATCAGCTGGC 488
QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
DB 489 AACACGGTGTCAAGTTCATCAATGAGGATTAATACCTTTGACAGGAATTTCTTGTCCAA 548
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 549 CTGTTAATGACGTGAGAGAGCTGCTTCAATCAGATCATTCAGGCTCACTCACTGAGAA 608

QY 141 SerHisGIYArGVaLaSnAsnValPheAspHisPheSerAspCysGluPheLeuAla160
Db 609 TGGCATGCAAGCTGTCAACATGCTTGTATCATCTTCGATTTGTAATTTTGGCTGCC 668
QY 161 LeuTYrAsnProPheGlyAsnPhelysProHisLeuGlnIlyLeuCysAspGlyIleAsn 180
Db 669 TTGTAAATACCTCTTTGGACCTTAAACCTTCAATTCCTTCAAAACCTTGTGTGATGCTCAAC 728
QY 181 LysMetLeuAspGluGluAsnIle 188
Db 729 AGAATCTAGATGAGGGAACATA 752
RESULT 25
AK097884 1766 bp mRNA linear PRI 30-JAN-2004
LOCUS AK097884
DEFINITION Homo sapiens cDNA FL40565 fis, clone RHXMU2004688, highly similar
to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK097884
VERSION AK097884.1 GI:21757783
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shibatani, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, M., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yoshida, M., Hoshino, T., Kusanagi, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togawa, S., Komai, F., Hara, R.,
Takuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Nomiyama, H., Sato, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Omori, Y., Kawabata, A., Hiki, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, A., and Sugano, S.
TITLE Complete sequencing and characterization of 21,243 full-length
human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702035
REFERENCE
AUTHORS 2
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Ota, T., Makamatsu, A., Ishii, S., Yamamoto, J.,
Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,
Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,
Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A.,
Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K.,
and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1766)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7

COMMENT
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Health Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..1766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RHXMU2004688"
/tissue_type="Chymus"
/clone_lib="RHXMU2"
/note="Cloning vector: pME18FL3"
ORIGIN
Alignment Scores:
Pred. No.: 4.25e-67 Length: 1766
Score: 856.50 Matches: 173
Percent Similarity: 92.06% Conservative: 1
Best Local Similarity: 91.53% Mismatches: 1
Query Match: 89.03% Indels: 14
DB: 9 Gaps: 1
US-10-627-571-2 (1-188) x AK097884 (1-1766)
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnIlyLysIleLeu 20
Db 176 ATGGCCACGATGCTTTTATTCCTCAAAACCTGGCCGCTTCAAGCAAAAGAGATCTTG 235
QY 21 GlyLysMetValSerLysSerIleAlaThrIleLeuIleAspAspThrSerGluVal 40
Db 236 GGTAAATAGTGTCCAAATCCATCCGCCACCACTTATAGACGACAAAGTAGTGCTG 295
QY 41 LeuAspGluLeuTYrArgValThrArgGluTYrThrGlnAsnIlyLeuAlaGluLys 60
Db 296 CTGGATGACCTTACAGA-----GAA 316
QY 61 Lys-IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTYrArgAsnAsnI 80
Db 317 GATCATCAAGAACCTTCATCAAGACAGTCACTCAAGCTGGCATTTTATAGAAATATATCA 376
QY 80 nPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAl 100
Db 377 GTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTAAAGAAAGATTCACTGCTGC 436
QY 100 aMetThrValValSerPheHisGlnValAspTYrThrPheAspAgaValLeuSerAr 120
Db 437 TATGACCGGTCGACATTTCCATCACTGATTTATACCTTTGACCGGAATGTGTATCCAG 496
QY 120 GLeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db 497 GCTGTAAATGATGAGAGAGATGCTGCACCAAAATCTTACGCGCACCTCACTGCCTCA 556
QY 140 sSerHisGIYArGVaLaSnAsnValPheAspHisPheSerAspCysGluPheLeuAla1 160
Db 557 GTCACATGAGACGGGTAAATATGATGCTTTGATCATTTTTCAGATTTGTAATTTTGGCTGC 616
QY 160 aLeuTYrAsnProPheGlyAsnPhelysProHisLeuGlnIlyLeuCysAspGlyIleAsn 180
Db 617 TTGTAAATACCTCTTTGGACCTTAAACCTTCAATTCCTTCAAAACCTTGTGTGATGCTCAAC 676
QY 180 nLysMetLeuAspGluGluAsnIle 188
Db 677 CAAATGTTGATGAGAGAACATA 701
RESULT 26
CR760636

LOCUS CR760636 1559 bp mRNA linear VRT 15-SEP-2004
 DEFINITION Xenopus tropicalis finished cDNA, clone TEG9053008.
 ACCESSION CR760636
 VERSION CR760636.1 GI:51966545
 KEYWORDS Xenopus tropicalis (Silurana tropicalis)
 SOURCE Xenopus tropicalis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus; Silurana.
 REFERENCE 1 (bases 1 to 1559)
 AUTHORS Amaya, E., Ashurst, J.L., Bonfield, J.K., Croning, M.D.R., Davies, R.M., Francis, M.D., Garrett, N., Gilchrist, M.J., Graham, D.V., McLaren, S.R., Papalopulu, N., Rogers, J., Smith, J.C., Taylor, R.G., Voigt, J. and Zorn, A.M.
 TITLE Direct Submission
 JOURNAL Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
 COMMENT Sanger Xenopus tropicalis EST/cDNA project.
 This sequence is from a Xenopus Gene Collection (XGC) library, from a library constructed by Aaron M. Zorn. cDNA was prepared from RNA extracted from eggs. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli XL1-blue.
 FEATURES
 Location/Qualifiers
 1..1559
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TEG9053008"
 /clone_1ib="XGC-egg"
 /dev_stage="egg"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.4e-66 Length: 1559
 Score: 844.50 Matches: 161
 Percent Similarity: 96.30% Conservative: 21
 Best Local Similarity: 85.19% Mismatches: 6
 Query Match: 87.79% Indels: 1
 DB: 5 Gaps: 1
 US-10-627-571-2 (1-188) x CR760636 (1-1559)
 QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlValGlnAlaGlnLysLysIleLeu 20
 Db 155 ATGGCAACAGATGTCCTTCAGTCCAAAATCTGGCTGTACAAAGCCAAAGAAAGATTCTT 214
 QY 21 GlyLysMetVal---SerLysSerIleAlaThrLeuIleAspAspThrSerSerGlu 39
 Db 215 GGGCAATGGCCCTCATCAAGATGATACATAGCAACATCTTATGACATCAAGCGGTGAA 274
 QY 40 ValLeuAspGluLeuLysrArGValThrArGValLysrThrGlnAsnLysLysGluAlaGlu 59
 Db 275 GTTCTAGATGAAGTGTACCATTTGACGAGGAGATACCCAGAGTAAGAAATCCGAA 334
 QY 60 LysLysIleLysAsnLeuIleLysrValIleLysLeuAlaIleLeuLysrArGAsnAsn 79
 Db 335 AAAGTTATAAAGAACTCATCAAAACCGTATCAAGCTGCACTCTGTACAGAAATAT 394
 QY 80 GlnPheAsnGlnAspGluLeuAlaLeuMetGlnLysPheLysLysValHisGlnLeu 99
 Db 395 CAGATTAAATGAGAAAGATGTGACCTCATGAGAGATTCMAAGCAAGGCCATCTCAGCTG 454
 QY 100 AlaMetThrValValSerPheHisGlnValAspLysrThrPheAspArGAsnValLleuSer 119
 Db 455 GCATATGCTGTGTGCTTATTCAGGTGGAATATACCTTTGACGAAACGAGCTTTCC 514
 QY 120 ArgLeuLeuAsnGlnLysrArGValMetLeuHisGlnIleIleGlnArGHisLeuThrAla 139
 Db 515 AAATTGTAAATGATGAGCGGAGTGTCTCATCAAGTATACAGGCCATCTCAGGGCA 574

QY 140 LysSerHisGlyArGValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 159
 Db 575 AAGTCTACGAGAGCTTATTAACGTTTGAACACTTTCAATATGTGATTTTGGCT 634
 QY 160 AlaLeuTyrAsnProPheGlyAsnPhelLysProHisLeuGlnLysLeuCysAspGlyIle 179
 Db 635 GCTCTTACATCCCTTTTGGACCTTTCAAAAGACACTTCAGAGACTCTGTGAAGAGTCT 694
 QY 180 AsnLysMetLeuAspGluGluAsnIle 188
 Db 695 AACAAATGCTGGACGAGCAATATATT 721
 RESULT 27
 LOCUS BC072904
 DEFINITION Xenopus laevis MG80354 protein, mRNA (cDNA clone MGC:80354 IMAGE:5074412), complete cds.
 ACCESSION BC072904
 VERSION BC072904.1 GI:49115085
 KEYWORDS MGC.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus; Xenopus.
 REFERENCE 1 (bases 1 to 1193)
 AUTHORS Klein, S.L., Struhsberg, R.L., Wagner, L., Pontius, J., Clifton, S.W. and Richardson, P.
 TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative
 JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
 PUBMED 12454917
 REFERENCE 2 (bases 1 to 1193)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, A.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 REFERENCE 3 (bases 1 to 1193)
 AUTHORS Klein, S. and Gerhard, D.S.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
 REMARK NIH-MGC Project
 COMMENT Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor David
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth

Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim Macdonald, Amara Maason, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liles Prambu, Paraneh Seedi, JR Santos, Angeliqne Scherch, Ursula Skalska, Duane Smalius, Jeff Stott, Miranda Tsai, George Yang, Jacque Schain, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRK Plate: 152 Row: J Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

```
1..1193
  /organism="Xenopus laevis"
  /mol_type="mRNA"
  /db_xref="taxon:8355"
  /clone="MGC:80354 IMAGE:5074412"
  /tissue_type="Ovary, Xenopus"
  /clone_lib="NICHD_XGC_Ov1"
  /lab_host="DH10B"
  /note="Vector: pCMV-SPORT6"
  1..1193
  /gene="MGC80354"
  /db_xref="LOCusID:443963"
  153..722
  /gene="MGC80354"
  /codon_start=1
  /product="MGC80354 protein"
  /protein_id="AAH72904.1"
  /db_xref="GI:49115086"
  /translation="MATTDFNSKNAVAQOKILGMASSKTIATSLIDTSGEVIDE
LTVQVREYQNKDAKTIKNIKTVIKIAVLVYRNQFBEIEGLKEKRVHQLAM
TVSPFYQVETPDRLVSKLNECRELLHQVORHLTAQSHGVNNVDFHFNCEFLA
ALNPFQPYKXHLORLQNGVNMKLEDENI"
```

CDS

gene

ORIGIN

Alignment Scores:

```
Pred. No.:      5,03e-66      Length:      1193
Score:          842.50      Matches:      160
Percent Similarity: 95.24%      Conservative: 20
Best Local Similarity: 84.66%      Mismatches: 8
Query Match:    87.58%      Indels:      1
DB:             5          Gaps:      1
```

US-10-627-571-2 (1-188) x BC072904 (1-1193)

```
QY      1 MetAlaThAspValPheAsnSerLyAsnLeuAlaValGlnAglInLySylLeu 20
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      153 ATGGCAACCGATATCTTCAATTCGAAAACCTGGCTGACAGCCGAAAGAGATTCTT 212
QY      21 GlyLysMetVal--SerLySerIleAlaThrThrLeuIleAspAspThrSerSerGlu 39
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      213 GCGAAGATGGCTTCATCAATATCATAAGCAACGTCCTGATGATCAACAGTGGCGAA 272
QY      40 ValLeuAspGluLeuLysrArgValThrArgGluTyrThrGlnAsnLysLysGluAglu 59
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      273 GTTCTAGATGAAGCTTACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 332
QY      60 LysLysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLysrArgAsnAn 79
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      333 AAAATTCACAAAGACTTCATCAAAAACAGTCATCAAGTGGCAGCTCTTGATCCGAAACAC 392
QY      80 GlnPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeu 99
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      393 CAGTTAAACGAGGAAGAGATGGTCTTATGAGAGAAAGTTAAAGGAAGGCCATCGCTG 452
QY      100 AlaMetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSer 119
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      453 GCATGACTGTGTGAGCTTCTATACGATGGAATATACCTTTGACCGAAACGCTTTCC 512
```

```
QY      120 ArgLeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAla 139
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      513 AAATTCTTAAAGATGACGAGGAGCTCTTCCACGATCATACAGCCATCTCACGGA 572
QY      140 LysSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAla 159
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      573 AAGTCTCAGCAGCGAGTGTATTAACGTTGTGACCACTTTCAAAATGTGAAATTTTACGT 632
QY      160 AlaMetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSer 179
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      633 GCTCTGTCAATCCCTTTGGACCTTACAAAACACCTTCAGAGACTGTGAACGAGATC 692
QY      180 AsnLysMetLeuAspGluGluAsnIle 188
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      693 AACAAATGCTGTGATGAACATATTT 719
```

RESULT 28

BC052765

LOCUS

DEFINITION

Danto rerio tumor necrosis factor, alpha-induced protein 8, mRNA (CDNA clone MGC:55331 IMAGE:2601021), complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Danto rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danto.

1 (bases 1 to 3986)

REFERENCE

AUTHORS

Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D.,

Alischul,S.F., Zeeberg,B., Buelow,K.H., Scheffer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldi,M.F., Cacaavant,T.L.,

Schaefer,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,

Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3986)

REFERENCE

JOURNAL

PUBMED

TITLE

AUTHORS

DIRECT SUBMISSION

JOURNAL

COMMENT

REMARK

CONTACT: MGC help desk

Email: cgabs-remail.nhl.nih.gov

Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami

CDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbiology.org>contact: amadn@systemsbiology.org

Amp Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>

Series: IRAC Plate: 101 Row: c Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gl: 41056158.
 Location/Qualifiers

FEATURES

source

```

1. 3986
/organism="Danio rerio"
/mol_type="mRNA"
/strain="AB"
/db_xref="taxon:7955"
/clone="MGC:55331 IMAGE:2601021"
/risue_type="whole body, adult, (one male and one female,
including unfertilized eggs)"
/clone_lib="Sugano Kawakami zebrafish DNA"
/lab_host="DH10B"
/notes="Vector: pME18S-PL3"
1. 3986
/gene="tnfaip8"
/notes="synonyms: MGC55331, zgc:55331"
/db_xref="LOCUSID:393303"
196. 756
/gene="tnfaip8"
/codon_start=1
/product="tumor necrosis factor, alpha-induced protein 8"
/protein_id="AAH52765.1"
/db_xref="GI:31127058"
/translation="MDSFSTKNLALQAKKMSKMTKTVANLFDTSSEVLDELVR
VTRKVRNKEAKQIKIKNLTKMVKGLVLRNQCFFNNEELATVERRKVHTLMTAV
SFYQIDFTFDRVMSNLNDCRELHQAIRHRLTASSHARINIVFHPADCDPLATLY
GPEYVRHGHQIKCEGVNKMLEDGNL"

```

ORIGIN

Alignment Scores:

Pred. No.: 1.37e-52 Length: 3986
 Score: 698.00 Matches: 125
 Percent Similarity: 86.49% Conservative: 35
 Best Local Similarity: 67.57% Mismatches: 25
 Query Match: 72.56% Indels: 0
 DB: Gaps: 5

US-10-627-571-2 (1-188) x BC052765 (1-3986)

```

QY 4 AspValPheasnserlysaenleuAlaValGlnAluGlnlysllyleuGlylyMet 23
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 199 GACTCGTTGACGACCAAGAGCTTATGAGACACCGACGAGGATCTGACGACATG 258
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 24 ValSerlyserlleaTnTrhLeuIleAspAspThrSerSerGluValIleuAspGlu 43
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 259 GCGACCAAGACAGTGGCCAACTCTTATGAGACACCGACGAGGATCTGACGAG 318
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 44 LeuTyArgValThrArgGluTyThrGlnAsnlysllysluAgluGlysllyle 63
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 319 CTGTACGAGTCAACCAAGAGTACACCGCACCGCAGAGGAGCCCGAAGATCATCA 378
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 64 AsnleuIlelyslThrValIlelyslleuAlIleleuTyArgAsnAsnGlnPheasnGln 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 379 AACCTCATCAAGATGGTGGTCAAGTTGGGCTCTTACCGCAACGGTCAATTCACAAC 438
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 84 AspGluLeuAlaleuMetGluTyPhelelysllyslValIleGlnleuAlaMetThrVal 103
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 439 GAGGAGTACCGTTAGTCGAGCGTTTCGTAAGAGGTGATCGCTGGCGATGACGCT 498
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 104 ValSerPheHisglValaAspTyThrPheAspArgAsnValleuSerArgleuLeuAsn 123
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 499 GTTAGCTTTCACAGATCGACTTCACTTTGATCGACGGCTCATGAGTATCTACTCAAT 558
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 124 GluCyArgGluMetleuHisglInleIleGlnArgHisleuThrAlaIlySerHisgl 143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 559 GACTGCGGTAACTTTCGACACGCGCATCAATCGGATCTTAAACGGGAATCTCAACGCC 618
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 144 ArgValaAsnAsnValPheAspHisPheSerAspCygluPheleuAlaIleuTyArgAsn 163
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 619 CGAATCAACATGCTTTCATATCTTTCGCGGATTTGACTTCTCGGACGCTTATACGGA 678

```

QY 164 ProPheGlyAsnPhelysProHisleuGlnlyleuCyArgAspGlyIleAsnlyMetleu 183
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 679 CTTTCAGAAAGTACCGCGCCACTTGCAGAAAGATCTGTGAAGAGATCAACAAATGCTG 738

QY 184 AspGluGluAsnIle 188
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 739 GACCAAGGCAATCTT 753

RESULT 29
 BX649252 205949 bp DNA linear VRT 06-FEB-2004
 BX649252/c
 LOCUS
 DEFINITION
 Zebrafish DNA sequence from clone DKEY-49M19 in linkage group 22,
 complete sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Danio rerio (zebrafish)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 205949)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Submitted (06-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 zfish-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Feb 6, 2004 this sequence version replaced gl:41392451.

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
 zebrafish pUC subclones occasionally display inconsistency over the
 length of mononucleotide A/T runs and conserved TA repeats. Where
 this is found the longest good quality representation will be
 submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat
 discovery system (Zhifeng Bao and Sean Bddy, submitted), and those
 beginning 'dir' were identified by Rick Waterman (Stephen Johnson
 lab, WashU). For further information see
 http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-49M19
 is from a Zebrafish BAC library
 VECTOR: pIndigoBAC-5.

FEATURES

source

```

1. 205949
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_lib="DKEY-49M19"
/clone_lib="Dantokay"

```

ORIGIN

Alignment Scores:

Pred. No.:	7,986-51	Length:	205949
Score:	698.00	Matches:	125
Percent Similarity:	86.49%	Conservative:	35
Best Local Similarity:	67.57%	Mismatches:	25
Query Match:	72.56%	Indels:	0
DB:	5	Gaps:	0

US-10-627-571-2 (1-188) x BX649252 (1-205949)

```

QY      4 AspValPheasnSerlysaenleuAlaValGlnAlaGlnlysllysleuGlyLysmet 23
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69523 GACTGCTTACAGCACCAAGAACTAGCCCTGCAGGCCGACAGAAAGTCTATGAGCAAGATG 69464
QY      24 ValSerLyseSerlleAlaThrThrleuIleAspAspThrSerSerGluValleuAspGlu 43
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69463 GCGACCCAGACAGTGGTCCCACTCTTTATAGACGACACGACGCGAGTACTGACGAG 69404
QY      44 LeuTyArgValThrArgGluTyThrGlnAsnlyslsGlnAlaGlnlysllysleu 63
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69403 CTGTACAGAGTCCCAAGAGGACACGCGACCGCAAGAGGCGCCAGAAAGATCATCAAA 69344
QY      64 AsnleuIlelyslThrValIlelyslsleuAlaIleLeuTyArgAsnAsnGlnPheasnGln 83
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69343 AACCTCATCAAGATGGTGGTCAAGTTGGGGGTCCTCTACCGCAAGGTCATTCAACAAC 69284
QY      84 AspGluLeuAlaLeuMetGluTybPheLyblyslsValIleGlnleuAlaMetThrVal 103
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69283 GAGGACCTGACGCTTATCGATCGAGCTTTTGTAAAGAGTCAATCGTCGATGACAGCT 69224
QY      104 ValSerPheHisGlnValAspTyThrThrPheAspAsnValleuSerArgleuLeuAsn 123
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69223 GTTAGCTTTACCAAGATCAACTCTTCTTGTATCGACGGCTCATGAGTAACTACTCAAT 69164
QY      124 GluCyArgGluMetleuHisGlnIleIleGlnArgHisleuThrAlaLysSerHisGly 143
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69163 GACTGCCGTCGAACTTCTGCACACGCGCATCAATCGGCATCTAAGCGGAATCTCAAGCC 69104
QY      144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyArgAsn 163
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69103 CCAATCAACATGTCCTTCAATCATCTTCCGCGATTTGACATCTCTCCGCGCATATCGGA 69044
QY      164 ProPheGlyAsnPheLyProHisleuGlnlyslsLeuCysAspGlyIleAsnlyslsMetleu 183
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      69043 CTTTCAAGAGTCAACCGCGGCCACTTGCAGAAAGATCTGTGAAGAGTCAACAGATGCTG 68984
QY      184 AspGluGluAsnIle 188
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      68983 GACGAAGCAATCTT 68969

RESULT 30
LOCUS   AJ720336                      1589 bp      mRNA       linear   VRT 30-SEP-2004
DEFINITION
Gallus gallus mRNA for hypothetical protein, clone 15114.
ACCESSION
AJ720336
VERSION
AJ720336.1 GI:53133331
KEYWORDS
ORF1.
SOURCE
Gallus gallus (chicken)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezzubov, Y., Zaim, J.,
Fiedler, P., Kuter, S., Blagodatski, A., Kostovska, D., Kotter, M.,
Plachy, J., Caminci, P., Hayashizaki, Y. and Buerstedde, J.M.
Full-length cDNAs from bursa lymphocytes to facilitate gene
function analysis
Unpublished
2 (bases 1 to 1589)
JOURNAL
Caldwell, R.B.
REFERENCE
AUTHORS
TITLE
Direct Submission

```

JOURNAL

Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuberg, GERMANY

FEATURES

source

location/Qualifiers

1..1589

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="CB"

/db_xref="taxon:9031"

/clone="15114"

/cell_type="lymphocyte"

/tissue_type="bursa"

/clone_lib="rikeni"

/dev_stage="2 weeks old"

1..1589

/locus_tag="RCMB04_15114"

<1..140

/locus_tag="RCMB04_15114"

141..701

/locus_tag="RCMB04_15114"

/note="ORF1"

/codon_start=1

/product="hypothetical protein"

/protein_id="CAG31995.1"

/db_xref="GI:53133332"

/translation="MDFTSTKNALAOAKKLSKMASTVANFIDTSSITLDELTYR

ATKEVTHNRKRAOKTIRKILIKYIMGLIYNGGSPPEILVMEFRKKVHTLMTAV

SFHQIDFTFDRRVSSVLTETCDLHQVSSHITKSRINHVNRHADYEFSLALY

GAEPYRTHLRKICGVNRMLEEDI"

702..>1589

/locus_tag="RCMB04_15114"

ORIGIN

3' UTR

Alignment Scores:

Pred. No.:	3,796-50	Length:	1589
Score:	666.00	Matches:	118
Percent Similarity:	84.32%	Conservative:	38
Best Local Similarity:	63.78%	Mismatches:	29
Query Match:	69.23%	Indels:	0
DB:	5	Gaps:	0

US-10-627-571-2 (1-188) x AJ720336 (1-1589)

```

QY      4 AspValPheasnSerlysaenleuAlaValGlnAlaGlnlysllysleuGlyLysmet 23
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      144 GACACCTTACAGCACCAAGAACTAGCCCTGCAGGCCGACAGAAAGTCTTGCAGCAAAATG 203
QY      24 ValSerLyseSerlleAlaThrThrleuIleAspAspThrSerSerGluValleuAspGlu 43
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      204 GCATCCAGACAGTGGCCAAATGCTTCAATGTATGACACAGCAGGAGATCTTGGATGAG 263
QY      44 LeuTyArgValThrArgGluTyThrGlnAsnlyslsGlnAlaGlnlysllysleu 63
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      264 CTCTACCGAGCCACCAAGAGATGACACTCAATCCCAAGAGAGCCGACCAAGATCAATCAAA 323
QY      64 AsnleuIlelyslThrValIlelyslsleuAlaIleLeuTyArgAsnAsnGlnPheasnGln 83
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      324 AACCTTATCAAGATTTGTATGAAGCTGGCGCTGCTGTACCGCAATGGCGATTCAGTCTCT 383
QY      84 AspGluLeuAlaLeuMetGluTybPheLyblyslsValIleGlnleuAlaMetThrVal 103
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      384 GAGGAGCTGCTGATGATGAGGAGGCTTCCGCAAGAAAGTGCACACCTTGGCCATGACAGCG 443
QY      104 ValSerPheHisGlnValAspTyThrThrPheAspAsnValleuSerArgleuLeuAsn 123
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      444 GTGAGCTTCCACCAAGATGACCTTACGTTTGAACCCGAGGCTGCTGCAAGTGTACAGTGCACA 503
QY      124 GluCyArgGluMetleuHisGlnIleIleGlnArgHisleuThrAlaLysSerHisGly 143
         |||.....|.....|.....|.....|.....|.....|.....|.....|
Db      504 GAGTCCGGGACCTCTCTGCACAGGCTGTCAGACGCGCACCTAAGCGCCAAAGTCAACATCC 563
QY      144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyArgAsn 163
         |||.....|.....|.....|.....|.....|.....|.....|.....|

```

Db 564 CGCATCAACACGCTTTAAACA CTTTGACAGACTGATGTTCCCTCGCCCTCTACGGG 623

Qy 164 ProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsnLysMetLeu 183

Db 624 CCGACCGAGCCCTACCGCACCACTGGAAGAGATCGAAGGTGGAACAAGATGCTG 683

Qy 184 AspGluGluAsnIle 188

Db 684 GAGAGAGACACACATA 698

RESULT 31

LOCUS CQ736656 1248 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 22590 from Patent WO02068579.

ACCESSION CQ736656

VERSION CQ736656.1 GI:42332642

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

TITLE Kites, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof

JOURNAL Patent: WO 02068579-A 22590 06-SEP-2002;

PE Corporation (NY) (US)

FEATURES

source 1. 1248

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 6,44e-49 Length: 1248

Score: 651.00 Matches: 114

Percent Similarity: 89.07% Conservative: 49

Best Local Similarity: 62.30% Mismatches: 20

Query Match: 67.67% Indels: 0

DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x CQ736656 (1-1248)

Qy 4 AspValPheAsnSerLysAsnLeuAlaValGlnAglInLysLysIleLeuGlyLysMet 23

Db 691 GATGTTTAAAGTTCACAAAGAGTCTTGCGCTTCAAGCCACAGAGAAGATTCTGACCAAAATA 750

Qy 24 ValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlu 43

Db 751 GCCAGCAAAACTGTGCCCAACATGTTGATTGATGACACCGACGAGATCTTGTATGAG 810

Qy 44 LeuTYArgValThrArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 63

Db 811 CTCTCAAAAGTCCCAAGAGCACACACACAAAGAGAGAACCCCAAGATCATGAAA 870

Qy 64 AsnLeuIleLysThrValIleLysLeuAlaIleLeuTYArgAsnAsnGlnPheAsnGln 83

Db 871 GACTTAATCAAGTGGCGATCAAAATCGGGATCTCTACCGGAACAACGATTACCAA 930

Qy 84 AspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAlaMetThrVal 103

Db 931 GAGGAGCTGGTATTTGTGAGAGAGTCCGGAGAAAGCTGAACACAGCCGCATGACCAT 990

Qy 104 ValSerPheHisGlnValAspTYThrPheAspArgAsnValLeuSerArgLeuLeuAsn 123

Db 991 GTCAAGCTTCTATAGAGTGAATACACTTCGATAGAGAACGTCTCTCAATCTCTGCAT 1050

Qy 124 GluCYArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGly 143

Db 1051 GAGTGCAGAGACCTGGTGCATGAACGTGTCAGCGGACCTGACGCCACGAGACCGAGG 1110

Qy 144 ArgValAsnAsnValPheAspHisPheSerAspCyGluPheLeuAlaIleLeuTYArg 163

Db 1111 CGCATCAACACGCTTTAAACA CTTTGACAGACTGATGTTCCCTCGCCCTCTACAGT 1170

Qy 164 ProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsnLysMetLeu 183

Db 1171 CTGATGAGAGACTGTAGGCCCAACCTCAAGAGATTGTGAAGATCATTAAGTTGCTA 1230

Qy 184 AspGluGlu 186

Db 1231 GATGAGAAA 1239

RESULT 32

LOCUS CQ841622 1986 bp DNA linear PAT 02-AUG-2004

DEFINITION Sequence 269 from Patent EP1440981.

ACCESSION CQ841622

VERSION CQ841622.1 GI:50893409

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Yamamoto, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,

TITLE Full-length human cdna

JOURNAL Patent: EP 1440981-A 269 28-JUL-2004;

Research Association for Biotechnology (JP)

FEATURES

source 1. 1986

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.04e-48 Length: 1986

Score: 651.00 Matches: 114

Percent Similarity: 89.07% Conservative: 49

Best Local Similarity: 62.30% Mismatches: 20

Query Match: 67.67% Indels: 0

DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x CQ841622 (1-1986)

Qy 4 AspValPheAsnSerLysAsnLeuAlaValGlnAglInLysLysIleLeuGlyLysMet 23

Db 152 GATGTTTAAAGTTCACAAAGAGTCTTGCGCTTCAAGCCACAGAGAAGATTCTGACCAAAATA 211

Qy 24 ValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlu 43

Db 212 GCCAGCAAAACTGTGCCCAACATGTTGATTGATGACACACGACGAGATCTTGTATGAG 271

Qy 44 LeuTYArgValThrArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 63

Db 272 CTCTCAAAAGTCCCAAGAGCACACACACAAAGAGAGAACCCCAAGATCATGAAA 331

Qy 64 AsnLeuIleLysThrValIleLysLeuAlaIleLeuTYArgAsnAsnGlnPheAsnGln 83

Db 332 GACTTAATCAAGTGGCGATCAAAATCGGGATCTCTACCGGAACAACGATTACCAA 391

Qy 84 AspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAlaMetThrVal 103

Db 392 GAGGAGCTGGTATTTGTGAGAGAGTCCGGAGAAAGCTGAACACAGCCGCATGACCAT 451

Qy 104 ValSerPheHisGlnValAspTYThrPheAspArgAsnValLeuSerArgLeuLeuAsn 123

Db 452 GTCAAGCTTCTATAGAGTGAATACACTTCGATAGAGAACGTCTCTCAATCTCTGCAT 511

Qy 124 GluCYArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGly 143

Db 512 GAGTGCAGAGACCTGGTGCATGAACGTGTCAGCGGACCTGACGCCACGAGACCGAGG 571

Best Local Similarity:	62.30%	Mismatches:	20
Query Match:	67.67%	Indels:	0
DB:	9	Gaps:	0
US-10-627-571-2 (1-188) x AK123281 (1-1986)			
QY	4	AspValpheaAsnSerLysAsnLeuValGlnAlaGlnLysSerIleLeuGlyLysMet	23
DB	152	GATGTTTATTGTTTAAAGAGCTTGGCCCTTCACAGCCAGAAAGATTCTGAGCAATA	211
QY	24	ValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlu	43
DB	212	GCCAGCAAAACGTGGGCCAACATGTTGATTGATGATACACAGCAGCGAGCTTGATGAG	271
QY	44	LeuTyrArgValThrArgGlyTyrThrGlnAsnLysGluAlaGlyLysIleLys	63
DB	272	CTTACAAAGATCCACCAAGACACACACACAAACAAAGAGAGCCCAAGATCATGAAA	331
QY	64	AsnLeuIleLysThrValIleLysLeuAlaIleLeuLysTyrArgAsnArgIleAsnGln	83
DB	332	GACTTAATCAAGGTGGCGATCAAAATGCGAGATCTTACCGGAAACACAGTTTAGCCA	391
QY	84	AspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrVal	103
DB	392	GAGAGAGCTGTTATTATGGAGAGAGTTCCGAGAGAGCTGAACACAGACCGCATGACATT	451
QY	104	ValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuAsn	123
DB	452	GTCAGCTTCTTATGAGGTGGAATACACCTTCGATGAGAACGTCTTCATCTTCCTCAT	511
QY	124	GluCysArgGluMetLeuHisGlnIleIleGlnAlaArgHisLeuThrAlaLysSerHisGly	143
DB	512	GAGGCAAGGACCTCGGTGATGAACTGAGTCAAGGAGGACCTGACGCCAGACCCAGCGG	571
QY	144	ArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAlaAlaLeuTyrAsn	163
DB	572	CGCATCAACACGCTTTAACCACTTTGCCGATGTGAGATTCTCTCACCCTTAAGT	631
QY	164	ProHegLysAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeu	183
DB	632	CTGAGTGGAGACTGTAGGCCCAACCTCAAGAGATTGTGAGAGATCAATTAAGTTCTA	691
QY	184	AspGluGlu 186	
DB	692	GATGAGAAA 700	
RESULT 34			
LOCUS	AC012678	188937 bp	DNA linear HTG 18-JUL-2000
DEFINITION	Homo sapiens chromosome 15 clone RP11-394B5, WORKING DRAFT		
SEQUENCE	SEQUENCE, 10 unordered pieces.		
ACCESSION	AC012678		
VERSION	AC012678.2	GI:6728953	
KEYWORDS	HTG, HTGS_PHASE1, HTGS_DRAFT.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	MullalyOta, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.		
AUTHORS	1 (bases 1 to 188937)		
JOURNAL	Bruno, D., Conn, L., Dela Rosa, M., Federpiel, N., Foreman, P.,		
AUTHORS	Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marache, R.,		
JOURNAL	Morehouse, A. J., Oefner, P., Palm, C. J., Ramirez, D., Wilhelm, J.,		
AUTHORS	Yu, S. and Davis, R. W.		
JOURNAL	Unpublished		
AUTHORS	2 (bases 1 to 188937)		
JOURNAL	Bruno, D., Conn, L., Dela Rosa, M., Federpiel, N., Foreman, P.,		
AUTHORS	Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marache, R.,		
JOURNAL	Morehouse, A. J., Oefner, P., Palm, C. J., Ramirez, D., Wilhelm, J.,		
AUTHORS	Yu, S. and Davis, R. W.		
JOURNAL	Direct Submission		
AUTHORS	Submitted (03-NOV-1999) DNA Sequencing and Technology Center,		
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304,		
AUTHORS	USA		

COMMENT

On Jan 21, 2000 this sequence version replaced gi:6223629.

----- Genome Center
Center: Stanford DNA Sequencing and Technology Development Center

Center code: SDSYDC
Web site: <http://sequence-www.stanford.edu/group/human/>
Contact: hum-info@sequence.stanford.edu
----- Project Information

Center project name: 698
Center clone name: RP11-394B5

----- Summary Statistics
Sequencing Vector: M13mp18; X02513
Chemistry: Dye primer; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 177081 bases at least Q40
Consensus quality: 185332 bases at least Q30
Consensus quality: 186766 bases at least Q20
Insert size: 185828; agarose-fp
Insert size: 188487; sum-of-contigs

Quality coverage: 5.3 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-contigs.
NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2170: contig of 2170 bp in length
2171 2220: gap of unknown length
2221 7218: contig of 4998 bp in length
7219 7268: gap of unknown length
7269 15922: contig of 8654 bp in length
15923 15972: gap of unknown length
15973 26869: contig of 10897 bp in length
26870 26919: gap of unknown length
26920 39746: contig of 12827 bp in length
39747 39796: gap of unknown length
39797 54669: contig of 14873 bp in length
54670 54719: gap of unknown length
54720 69897: contig of 15178 bp in length
69898 69947: gap of unknown length
69948 93475: contig of 23528 bp in length
93476 93525: gap of unknown length
93526 138958: contig of 45433 bp in length
138959 139008: gap of unknown length
139009 188937: contig of 49929 bp in length.

FEATURES

source

1.188937
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/clone="RP11-394B5"
/clone_11b="RPCT human BAC library 11"

ORIGIN

Alignment Scores:

Pred. No.: 1.14e-46 Length: 188937
Score: 651.00 Matches: 114
Percent Similarity: 89.07% Conservative: 49
Best Local Similarity: 62.30% Mismatches: 20
Query Match: 67.67% Indels: 0
DB: 2 Gaps: 0

US-10-627-571-2 (1-188) x AC012678 (1-188937)

Qy 4 AspValPheAsnSerIleAlaValAlaGlnAlaGlnIleLysIleLeuGlyLysMet 23
DB 51694 GATGTTTATTAGTTCAAGAGTCTTGGCTTCAAGCCCAAGAGAGATTCTGAGCAAAATA 51753

Qy 24 ValSerIleSerIleAlaThrThrIleuIleAspAspThrSerGluValIleuAspGlu 43
Db 51754 GCCAGAAACGTGGCCCAACATGTTGATTGAGACCAAGCAGCATCTTGTATGAG 51813
Qy 44 LeuTyArgValThrArgGluTyThrGlnAsnIleValGluAlaGluLysIleLys 63
Db 51814 CTCTCAAAAGTCAACCAAGACACACACACACACACACACACACACACACACATGATGAA 51873
Qy 64 AsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnGlnPheAsnGln 83
Db 51874 GACTTATCAAGGTGGCGATCAAAATCGGATCTCTTACCGGAACACAGTTTACCA 51933
Qy 84 AspGluLeuAlaLeuMetGluLysPheLysIleValHisGlnIleuAlaMetThrVal 103
Db 51934 GAGGAGCTGTTATCTGTGAAGAGTTCCGGAAGACCTGAACCAAGCCCATGACCAT 51993
Qy 104 ValSerPheHisGlnIleValAspTyThrPheAspArgAsnValIleuSerArgLeuAsn 123
Db 51994 GTCAAGTTTATAGAGTGGAATACACCTTCGATAGGAACGTCTCTCCATCTCTGCA 52053
Qy 124 GluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGly 143
Db 52054 GAGTCAAGACACTGGTGCATGACATGATGACGCGCACCTGACCCCAAGCCACG 52113
Qy 144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeuTyAsn 163
Db 52114 CCATCAACACACACTCTTACCACTTGCAGATGAGATCTCTCCACCTCTATAGT 52173
Qy 164 ProPheGluAsnPheLysProHisIleuGlnIleLysLeuCyAspGlyIleAsnIleMetLeu 183
Db 52174 CTGATGAGACTGTTAGGCCCACTCAAGAGATTGTGAAGATCAATATGTTGCTA 52233
Qy 184 AspGluGlu 186
Db 52234 GATGAGAA 52242

RESULT 35
AC073964/c 189796 bp DNA linear PRI 23-FEB-2001
LOCUS Homo sapiens chromosome 15 clone RP11-394B5 map 15q21.2, complete
DEFINITION sequence.

AC073964 GI:13112225
VERSION AC073964.3
KEYWORDS HTG.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)
Qy 1 (bases 1 to 189796)
Db 1 (bases 1 to 189796)

Center code: UMWSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowen@systemsbiology.org
Drafting center: SDSTDC

----- Summary Statistics -----
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-terminator Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399

Note: Date from overlapping BACs AC022407 (drafting center: UMWSC)
and AC012169 (drafting center: UMWSC) were added for finishing

FEATURES

source

1..189796

Location/Qualifiers

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q21.2"
/clone_id="RP11-394B5"

/note="Data from overlapping BACs RP11-105D1 and
RP11-522G20 were added and the consensus sequence was
determined from RP11-394B5 to the extent possible"

misc_feature

1..15788

misc_feature

17935..17940

misc_feature

55786..56417

misc_feature

/note="sequence data generated from subcloned PCR product
only"

misc_feature

55786..56417

misc_feature

65115..65120

misc_feature

104690..104715

misc_feature

105973..189796

misc_feature

/note="Overlap with RP11-522G20 AC012169"

misc_feature

161495..161500

/note="low quality data"

ORIGIN

Alignment Scores:

Pred. No.:

1.158-46

Length: 189796

Score:

651.00

Matches: 114

Percent Similarity:

89.07%

Conservative: 49

Best Local Similarity:

62.30%

Mismatches: 20

Query Match:

67.67%

Indels: 0

Gaps:

9

DB:

0

US-10-627-571-2 (1-188) x AC073964 (1-189796)

QY

4 AspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlySylIleuGlyIysMet 23

DB

68379 GATGTTTAACTTCAAGAGAGTTCGCTCAAGCCGAGAGAAATTCGACAAATTA 68320

QY

24 ValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValIleuAspGlu 43

DB

68319 GCCACAAAGTGTGGCCAACTGTGTGATGTGACACGACGACGAGATCTTGTATGAG 68260

QY

44 LeuTYRAGValThrArgIlyuThrGlnAenIlySylValIleuIlySylIleuSyl 63

DB

68229 CTTCACAAAGTCAACAAAGACACACAAAGAAAGAACCCACAAAGATCTTGACCA 68200

QY

64 AsnLeuIlySylThrValIleuSylValIleuIlySylValIleuIlySylValIleuIlySyl 83

DB

68139 GAGTAAATCAAGGTGGCATCAAAATCGGATCTCTACCGGACACCAAGATTTGACCA 68140

QY

84 AspGluLeuAlaLeuMetGluIlySylPheIlySylValIleuIlySylValIleuIlySylValIleuIlySyl 103

DB

68139 GAGGAGCTGTATATGTGTGAGAAAGTTCCGGAAGAGCTGAACCAAGCCGATGACCAT 68080

QY 104 ValSerPheHisGlnValAspTYRThrPheAspArgAsnValIleuSerArgLeuLeuAsn 123

DB 68079 GTCAAGCTTCTATAGAGGTGGAAATCACTTCATAGGAACGCTCTCCAACTCTCAT 68020

QY 124 GluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIlySerHisGly 143

DB 68019 GAGTGCAAGACCTGTGTGATGAACTGTGACGAGGACCTGACGCCACCAAGACCAAGCGG 67960

QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysArgGluPheLeuAlaIleuTYRAsn 163

DB 67959 CGCATCAACCAAGCTTCTTAACTTGCATGTGAGATGAGTCTCTCCACCTCTAATAGT 67900

QY 164 ProPheGluAsnPheIlySylProHisLeuGlnIlySylLeuCysAspGlyIleAsnIlyMetLeu 183

DB 67899 CTGATGAGACCTGTAGGCCCAACCTCAAGAGATTTGTGAAGAAATCAATAGTTGCTA 67840

QY 184 AspGluGlu 186

DB 67839 GATGAGAA 67831

RESULT 36

AC115187_1/c

WPCOMMENT

Sequence split into 5 fragments LOCUS AC115187 Accession AC115187

Fragment Name Begin End

AC115187_1 1 110000

AC115187_2 200001 210000

AC115187_3 200001 310000

AC115187_4 300001 410000

AC115187_5 400001 420106

Continuation (2 of 5) of AC115187 from base 100001 (AC115187 Rattus norvegicus clone CH2

US-10-627-571-2 (1-188) x AC115187_1 (1-110000)

QY 5 ValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlySylIleuGlyIysMetVal 24

DB 32867 GTGTTTAACTTCAAGAGAGTTCGCTCAAGCCGAGAGAAATTCGACAAATTA 32808

QY 25 SerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValIleuAspGlu 44

DB 32807 AGCAAAAGTGTGGCCAACTGTGTGATGTGACACGACGACGAGATCTTGTATGAGCTG 32748

QY 45 TyrArgValThrArgIlyuThrGlnAenIlySylValIleuIlySylIleuSyl 64

DB 32747 TACAAGTCAACGAAAGACCAACCAAGAAAGAGGCCCAAGATCATCAAGAT 32688

QY 65 LeuIlySylThrValIleuSylValIleuIlySylValIleuIlySylValIleuIlySyl 84

DB 32687 GTAATTAAGGTGGGATCAAAATGTGTTCTTACCGGAAACAACATCTGACAGAG 32628

QY 85 GluLeuAlaLeuMetGluIlySylPheIlySylValIleuIlySylValIleuIlySylValIleuIlySyl 104

DB 32627 GAGGTGTGATGTGGAGAAATCTCGGAGAAATCAAGACCAAGATCATCAAGATGCTC 32568

QY 105 SerPheHisGlnValAspTYRThrPheAspArgAsnValIleuSerArgLeuLeuAsn 124

DB 32567 AGCTTCAACCAAGGTGAGTAACTTGTATCGAATGTGCTCTTAAACTTCTCATGAG 32508

QY 125 CysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIlySerHisGlyArg 144

DB 32507 TGCAGGACCTGTGCTCATGAATCTGTACGAGCACTTGAAGCCCAAGAACCAAGAGAGC 32448

QY 145 ValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTYRAsnPro 164

DB 32448

DB 32447 ATAAACCATGTCTTCAACACCTTCGCTGATGTGAGATTCCTTTCCACTCTCACTGACGCG 32388

QY 165 PheGlyAsnPhenylProHISleuGlnIlySleuCyAspGlyIleAsnIlySmetLeuAap 184
 DB 32387 CACGGAACCTGACGCCCATCTCAAGAGATTTCGAAGAGATCAATTAATTTGATGAT 32328
 QY 185 GIUGlu 186
 DB 32327 GACAAA 32322

RESULT 37
 AC115505/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-77C4, *** SEQUENCING IN PROGRESS ***
 9 unordered pieces.
 AC115505
 AC115505.4 GI:25188467
 HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS 1 (bases 1 to 256285)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D., Anyelebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F., Bivak, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Cencer, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harrey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A., Johnson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, J., Kovar, C., Kowitz, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuwa, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawliny, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Natr, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokelelemah, O., Okwomou, G., Olarnpunaagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajls, D., Snead, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J., Steimle, M., Strong, R., Sutron, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umam, K., Valdes, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Direct Submission
 Unpublished
 2 (bases 1 to 256285)
 Worley, K.C.
 Direct Submission
 Submitted (20-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 256285)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Nov 23, 2002 this sequence version replaced gi:23269911.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSWY
 Center clone name: CH230-77C4
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 199255 bases at least Q40
 Consensus quality: 203424 bases at least Q30
 Consensus quality: 206379 bases at least Q20
 Estimated insert size: 195822; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 24557: contig of 24557 bp in length
 24558 24657: gap of unknown length
 24658 201810: contig of 17715 bp in length
 201811 201910: gap of unknown length
 201911 214313: contig of 12403 bp in length
 214314 214413: gap of unknown length
 214414 247000: contig of 32587 bp in length
 247001 247101: gap of unknown length
 247102 248513: contig of 1413 bp in length
 248514 249613: gap of unknown length
 249614 249866: contig of 1373 bp in length
 249867 250086: gap of unknown length
 250087 251489: contig of 1403 bp in length
 251490 251589: gap of unknown length
 251590 252936: contig of 1347 bp in length
 252937 253036: gap of unknown length
 253037 256285: contig of 3249 bp in length.

Location/Qualifiers
 1..256285
 /organism="Rattus norvegicus"

```

/mol type="genomic DNA"
/db xref="taxon:10116"
/clone="CH230-77C4"
misc_feature
1..1142
/note="wgs_contig"
misc_feature
4086..6215
/note="wgs_contig"
misc_feature
24658..25854
/note="wgs_contig"
misc_feature
228541..229886
/note="wgs_contig"
misc_feature
245270..247000
/note="wgs_contig"
ORIGIN
Alignment Scores:
Pred. No.: 6.3e-45 Length: 256285
Score: 633..00 Matches: 109
Percent Similarity: 87.91% Conservatave: 51
Best Local Similarity: 59.89% Mismatches: 22
Query Match: 65.80% Indels: 0
DB: 2 Gaps: 0
US-10-627-571-2 (1-188) x AC115505 (1-256285)
Oy 5 ValPheAenSerLySAnLeuAlaGlnAlaGlnLySLeuGlyLeuGlyLeuVal
Db 120592 GTGTTACTCTCAAGAGTCTGCCCTTCAAGCCCGAAGAGATCTTGAAGCAAGTACG
Oy 25 SerLySeriLeAlaThrThrLeuLeaSpAphrThrSerSerGlyValLeuApsGluLeu
Db 120532 AGCAAAACCTGGCGCAACATGCTGATGATGACACACAGAGGATCTTGTGATGAGCTG
Oy 45 TyrArgValITrArgGluTrThrGlnInuLySylsValAlaGlnLySylsLeuVal
Db 120472 TACAAAGTCACGGAAGAACACACCCACAAAGAGAGGCCACAGATCATGAAGAT
Oy 65 LeuileYsThrValileYsLeuAlaleuYrAgsAnSngInPheAnSngInAsp
Db 120412 GTANNAAGGCGGCATCAAAATTGCTATTCCTACCGGAACAAACAGTTCAAGTCAAGG
Oy 85 GluLeuAlaLeuMetGluLySphelyslyslsValhSgInLeuAlaMetThrValVal
Db 120352 GAGGTTCGATTTGGAGAAACTCCGGAAGAAAGCAAGACCGCATGACCATGTC
Oy 105 SerPhehSgInValAspTrThrPheAspAgsAnValLeuSerArgLeuLeuAnSgu
Db 120292 AGCTTCTACGAAGTGAAGTATACCTTTATACGAATGTGCTCTTAACTTCTGATGAG
Oy 125 CysArgGluMetLeuHISGlnlelleGlnArgHISleuThrAlaLySerHISGlyArg
Db 120232 TGCAGAGACCTGGTCAATGACTGTACGACGACCTTGACGCCCAAGAACCCAGGAGCC
Oy 145 ValhAnAnValPheAspHISpSerAspCysGluPheLeuAlaAlaLeuYrAnSPro
Db 120172 ATAAACCATGTCCTTCAACCATCTTCAATGATGGAATTCCTTCAACCTTCAAGTCG
Oy 165 PheGlyAenPheLyRchISleuGlnLySleuCysArgGlylleAsnLyMetLeuAsp
Db 120112 CACGGAACCTGACAGCCCAATCTCAAGAGGATTTGCGAAGGAATCAATTAATTTTAAT
Oy 185 GluGlu 186
Db 120052 GACAA 120047

```

SOURCE
ORGANISM
Rattus norvegicus (Norway rat)

REFERENCE
AUTHORS
Muzny D,Marie., Meeker,M,lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alebrooks,S., Amin,A., Angiano,D., Anyalbechi,V., Aoyagi,A., Ayodeji,M., Bacc,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnesread,M., Benahmed,F., Bismail,K., Blair,U., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Dudin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jolivet,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Kapathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Koyar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,D., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Loresuhewa,L., Louissege,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Matlinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munnidasa,M., Murphy,M., Natr,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakoilemech,O., Okwoum,G., Olarpunsaqoon,A., Pal,S., Parke,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannoch,C., Plopper,P., Poindexter,A., Popovic,D., Primus,E., Pu,L.L., Puzo,M., Quiroz,J., Rachlin,B., Reeves,K., Regler,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Slisson,I., Silter,C.D., Snajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabot,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasaana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Wooden,H., Woley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausem,A., Weiser,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 273225)
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 273225)
Direct Submission
Rat Genome Sequencing Consortium.
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:22364293.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described

TITLE
JOURNAL
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seman, S., Severy, P., Spencer, B., Stange, Thoman, N., Stojanovic, N.,
Strahan, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.
Submitted (21-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164258)
Birtten, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barina, N., Bastien, Y., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Maddison, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menais, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thoman, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.
Submitted (24-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 24, 2003 this sequence version replaced gi:28261544.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: 174.O.15
Center clone name: 174.O.15

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162606 bases at least Q40
Consensus quality: 162965 bases at least Q30
Consensus quality: 163183 bases at least Q20
Insert size: 164000; agarose-IP
Insert size: 163558; sum-of-coverage
Quality coverage: 11.0 in Q20 bases; agarose-IP
Quality coverage: 11.0 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 29707: contig of 29707 bp in length
*
* 29708 29807: gap of 100 bp
* 29808 32340: contig of 2533 bp in length
* 32341 32440: gap of 100 bp
* 32441 37689: contig of 5249 bp in length
* 37690 37789: gap of 100 bp
* 37790 104529: contig of 66740 bp in length
* 104530 104629: gap of 100 bp

FEATURES
source
1..164258
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-174015"
/clone_id="RP24-174015"
1..29707
/note="assembly_fragment"
clone_end:SP6
vector_side:left
misc_feature
29808..32340
/note="assembly_fragment"
32441..37689
/note="assembly_fragment"
37790..104529
/note="assembly_fragment"
104630..115026
/note="assembly_fragment"
115127..131677
/note="assembly_fragment"
131778..158793
/note="assembly_fragment"
158894..164258
/note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN
Alignment Scores:
Pred. No.: 1.68e-44 Length: 164258
Score: 626.00 Matches: 109
Percent Similarity: 86.26% Conservative: 48
Best Local Similarity: 59.89% Mismatches: 25
Query Match: 65.07% Indels: 0
DB: 2 Gaps: 0
US-10-627-571-2 (1-168) x AC118938 (1-164258)
OY 5 ValpheansSerIysAsnLeuAlaValGlnAlaGlnIysIysIleLeuGlyLysMetVal 24
DB 87991 GTGTTAGTTCAGAAATCTGCGCTTCAAGCCAGAAAGATCTGAGCAAGATAGCC 87932
OY 25 SerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeu 44
DB 87931 AGCAAACTGGGCAACATGCTGATTGATACACGACGTGATCTTGGATGAGCTG 87872
OY 45 TyrArgValThrArgGluYrThrGlnAsnIysGlnIaGluIysIleLysAsn 64
DB 87871 TACAAAGTCACGAATAACACACCCACACAAAGAGAGGCCACCAAGATATATAAGAT 87812
OY 65 LeuIleYrThrValIleIysLeuAlaIleLeuYrArgAsnAngInpAsnGlnAsp 84
DB 87811 GCTATCAAGGCGCAATCAAAATGCGATCTTACCGGAACAAACAGTTCAAGAG 87752
OY 85 GluLeuAlaLeuMetGluYrPheLysLysValIleGlnLeuAlaMetThrValVal 104
DB 87751 GAGGTTATTAATTGGAGAACTCCGAGAAAGACTGAACAGACTGGATGACATGGTC 87692
OY 105 SerPheIleGlnValAlaPyrYrThrPheAspArgAsnValLeuSerArgLeuAsnGlu 124
DB 87691 AGCTTTCACGAAGTACGATACCTTCTGATACGAATGCTCTTAAAGCTTTCATGAG 87632
OY 125 CysArgGluMetLeuIleGlnIleIleGlnArgIleLeuThrAlaLysSerHisGlyArg 144
DB 87631 TGCAAGACCTGTAACATGAAGTCAAGGAGCACTTGACACCCAGAAACCATGAGAGC 87572

QY 145 ValAsnAsnValPheAspHisPheSerAspCysGluThelenuAlaAlaLeuTyrAsnPro 164
 DB 87571 ATAAACGATGCTTCAACCACTTTGATGTGAATGAACTCTTCACTCTTATAGTCCG 87512
 QY 165 PheGlyAsnPhenylPProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeuAsp 184
 DB 87511 CATGGAACCTGCAGGCCCAATCTCAAGAGATTGCCAGAGATCAATAAATGTGTAGAC 87452
 QY 185 GluGlu 186
 DB 87451 GACAG 87446

RESULT 40
 AC107848
 LOCUS
 DEFINITION Mus musculus chromosome 9 clone RP23-321A19 map 9, *** SEQUENCING
 IN PROGRESS ***, 8 unordered pieces.
 AC107848
 VERSION AC107848.5 GI:45589920
 KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 195574)
 Mus musculus chromosome 9, clone RP23-321A19
 2 (bases 1 to 195574)
 Unpublished

REFERENCE
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B.,
 Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,
 Choquel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kellis, C., LaRocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
 MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T.,
 Mienna, P., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Sudramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 195574)
 Birren, B., Nusbaum, C., Lander, E., Abouelellail, A., Allen, N.,
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choquel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., Dearrellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kellis, C., Lander, E., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Menus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N.,
 Reichpka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,

JOURNAL

Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (22-MAR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 22, 2004 this sequence version replaced gi:44886703.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L20615
 Center clone name: 321_A_19

NOTE: This is a 'working draft' sequence. It currently
 consists of 8 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 7878: contig of 7878 bp in length
 1 7879: gap of 100 bp
 1 7979: contig of 2052 bp in length
 1 28231: gap of 100 bp
 1 28330: gap of 100 bp
 1 28331: gap of 100 bp
 1 59488: contig of 3118 bp in length
 1 59489: gap of 100 bp
 1 59589: gap of 100 bp
 1 100956: contig of 41368 bp in length
 1 100957: gap of 100 bp
 1 101057: gap of 100 bp
 1 103534: contig of 2478 bp in length
 1 10355: gap of 100 bp
 1 10355: gap of 100 bp
 1 103635: gap of 100 bp
 1 142683: contig of 39049 bp in length
 1 142684: gap of 100 bp
 1 142784: gap of 100 bp
 1 154264: contig of 11480 bp in length
 1 154264: gap of 100 bp
 1 154364: gap of 100 bp
 1 154364: contig of 41211 bp in length.

FEATURES

source
 1. 195574
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="9"
 /map="9"
 /clone="RP23-321A19"
 /clone_1lb="RPCT-23 Female Mouse BAC"

ORIGIN

Alignment Scores:
 Pred. No.: 2.01e-44 Length: 195574
 Score: 626.00 Matches: 109
 Percent Similarity: 86.26% Conservative: 48
 Best Local Similarity: 59.89% Mismatches: 25
 Query Match: 65.07% Indels: 0
 DB: 2 Gaps: 0

US-10-627-571-2 (1-188) x AC107848 (1-195574)

QY 5 ValPAsnSerIleAlaThrLeuAlaValGlnAlaGlnLysLysIleLeuGlnLysMetVal 24
 DB 160005 GGTATTAGTTCAGAGATCTCGCCCTTCAAGCCAGAGAGATCTTGAACAGATAGCC 160064
 QY 25 SerIleSerIleAlaThrLeuAlaValGlnAlaGlnLysLysIleLeuGlnLysMetVal 44
 DB 160065 AGCAAACTGTGCAGCAATCTGATGATGACACAGCAGTGATGATCTTGTATGAGCTG 160124
 QY 45 TyrArgValThrArgGluTyrThrGlnAlaValGlnAlaGlnLysLysIleLeuGlnLysMetVal 64

QY	Db	160125	TACGAAGTCACGAAATATACACACCCACACAGAGAGAGGCCACCAAGATCATGAAGAT	160184
QY	65	LeuilelyethrValillelyleuAlaleleuTyArAgaAnaInpheaengInaap	84	
Db	160185	GCTATCAAGGTGGCAATCAAAATTTGGCATTTCTTACCGGAACAAACAGTTCAAGTCAAGAG	160244	
QY	85	GIUleuAlaleuMeGIuIuysPheylslyblyValHiGIInleuAlamethrValVal	104	
Db	160245	GAGGTATATATTTGGAGAAACTCCGGAAGAACTGAAACCAAGCTCGATGACATAGTTC	160304	
QY	105	SePheHieSglInValAspTyThrPheAPArGaSnValIleuSerArgIleuLeuAnGIu	124	
Db	160305	AGCTTCTACGAAGTAGAGTACATCTTTGATACGAATGTCTCTCTTAAGCTTCTGCATGAG	160364	
QY	125	CyArGgluMetLeuHieSglInlelleGIaArgHsleuThAlAlaIysSerHieSglArG	144	
Db	160365	TGCAAGACAGCTGTACATGAATCTGGTACAGCGACATTTGACACCCAGAACCATGAGCG	160424	
QY	145	ValAAnaArValPheAspHisPheSerArPySglInPheuAlAlaleuTyArAnPro	164	
Db	160425	ATAAACCAAGTCTTCAACACACTTGTGTATGTGAATTCCTTCCACTCTCATATGATCGG	160484	
QY	165	PheGIyAnSPheIyProHieSglInIyIleuGIyAsPglYrIleAnaIyMetLeuAap	184	
Db	160485	CATGAAACTGCAGGCCCATCTCAAGAGATTTGCCAAGAAATCAATAAATTTGTTAGAC	160544	
QY	185	GIuGIu	186	
Db	160545	GACCAAG	160550	
RESULT 41				
BC076797		1938 bp	mRNA	linear
LOCUS		Xenopus laevis MGC83729 protein, mRNA (cDNA clone MGC:83729		
DEFINITION		IMAGE:6639647), complete cds.		
ACCESSION		BC076797		
VERSION		BC076797.1	GI:48999742	
KEYWORDS		MGC.		
SOURCE		Xenopus laevis (African clawed frog)		
ORGANISM		Xenopus laevis		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;		
		Xenopodinae; Xenopus; Xenopus.		
		1 (bases 1 to 1938)		
		Klein,S.L., Stransberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.		
		and Richardson,P.		
		Genetic and genomic tools for Xenopus research: The NIH Xenopus		
		initiative		
		Dev. Dyn. 225 (4), 384-391 (2002)		
JOURNAL				
PUBMED		12454917		
REFERENCE		2 (bases 1 to 1938)		
AUTHORS		Stransberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,		
		Klausner,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schlier,G.D.,		
		Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhac,N.K.,		
		Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,Y., Hsieh,F.,		
		Diatchenko,K.L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,		
		Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,		
		Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,		
		Carrinoni,P., Prange,C., Raha,S.S., Loguelfano,N.A., Peters,G.J.,		
		Abrahamson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,		
		McKernan,K.D., Malek,J.A., Gunaratne,P.H., Richards,S.,		
		Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huiyk,S.W.,		
		Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,		
		Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,		
		Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,		
		Boulfard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,		
		Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,		
		Butterfield,Y.S., Krzywinski,I., Sklisa,U., Smalish,D.E.,		
		Scherer,A., Schein,J.B., Jones,S.J. and Marra,M.A.		
		Generation and initial analysis of more than 15,000 full-length		
		human and mouse cDNA sequences		
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED		12477932		

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 1938)
Klein,S. and Gerhard,D.S.
Direct Submission
Submitted (06-JUL-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-XGC Project
Contact: XGC help desk
Email: cgabs-xemail.nih.gov
Tissue Procurement: Martha Rabbert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

REMARK
COMMENT

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liso,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Telika Olson, Diana Palmquist, Anca Petrescu, Anna Lissa Prabh,
Parvaneh Saeedi, J.R Santos, Angeliq Scherch, Ursula Skalska,
Diane Smilun, Jeff Strotz, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES
source

Clone distribution: WGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov
Series: IRAX Plate: 158 Row: d Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers

1. 1938
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="WGC:83729 IMAGE:6639647"
/tissue_type="Oocytes"
/clone_id="NICHD_XGC_OO1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"

gene

1. 1938
/gene="WGC83729"
/db_xref="LOCUSID:446389"
94. 702
/gene="WGC83729"
/codon_start=1
/product="WGC83729 protein"
/protein_id="AAH6797.1"
/db_xref="GI:49889743"
/db_xref="LOCusid:446389"
/translation="MDPDSGDLSEGLSPGPBOFSKFAVVOAKKLISMKATPMAN
MLDDTSSEIFDELFKYTKRYGNKKKAHHAYLKDYAVKVGKILRNPKQSLSEELJEI
VENRRKLNQTCMTAASFVEYTFPDNVLSGVLLHECKTLHLVPHRLVPKSHRID
RVNHFAVDVFPLALYSLEGNYPYLKRICEGVNKLLDERVL"

CDS

Alignment Scores:

Pred. No.: 8.55e-45 Length: 1938
Score: 607.00 Matches: 106
Percent Similarity: 83.89% Conservative: 45
Best Local Similarity: 58.89% Mismatches: 29
Query Match: 63.10% Indels: 0
DB: 5 Gaps: 0

US-10-627-571-2 (1-188) x BC076797 (1-1938)

Oy PheanSerLySaDeuAlaValAGlnALylLySrlLeuengLYlYSmetValSer 25
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db TTTAGCTCAAGACCTTGCTGTTCACGCCAAGAAGATTCTCAGTAAGATGGCACAA 210
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy LySeerlleAlaThrThrleuileaspaSpThrSerSerGluValleueapGluLeuTyrr 45
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::


```

Db      211  AAGACATGGCGCAATGCTTATTTGATGACACAGCATGTAATATTGAGACTCTTT 270
Qy      46  ArgValIthArggluYrThrGlnAsnLysLysGluAgluYsLysIleYsAsnLys 65
Db      271  AAGATTAACCAAGAGATATGTAAAGACAGAAAGAGCCCAAGATCTTGAAGACTT 330
Qy      66  IleYsThrValIleYsLysLysLysLysLysLysLysLysLysLysLysLysLys 85
Db      331  GTAAAGGTGCGAATTAAGTGGCATTTCTGTACAGAAATTAACAGTTGCTGGAAGAA 390
Qy      86  LeuAlaLeuMetGluYsPheYsLysLysLysLysLysLysLysLysLysLysLys 105
Db      391  CTGCAATATGTCGAAATCTTCAAGAAAATACTGAATCAAGACCTGCAATGCACTGACT 450
Qy      106  PheHisGlnValAspYrThrPheAspArgAsnValLeuSerArgLeuAsnGluCys 125
Db      451  TTCTTTAGAGTTGAGTACACGCTTGTATTAATGTGCTTCGCGAGCTTGCAAGATGT 510
Qy      126  ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal 145
Db      511  AAAACTCTTCTTCATGATTTGTCACGCGCATCTTAACCAAAATCCACAGTCGCATC 570
Qy      146  AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuYrAsnProPhe 165
Db      571  GACAGGCTTTTATCATTTTGTGCTGATGTGAATTCCTTACTGCTCTTATAGCTTGAA 630
Qy      166  GlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeuAspGlu 185
Db      631  GGAATATTACAGAGCGCTACCTCAAAAGATTTGTGAAGGGGTCAATAAATTACTTGAGAA 690

RESULT 42
BC053167      1813 bp      mRNA      linear      VRT 30-JUN-2004
LOCUS         Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907),
DEFINITION    complete cds.
ACCESSION     BC053167
VERSION       BC053167.1 GI:31418810
KEYWORDS      MGC.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Danio rerio
REFERENCE     1 (bases 1 to 1813)
AUTHORS       Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haiech, F.,
              Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stachert, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
              Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
              McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
              Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultyk, S.W.,
              Vallaloun, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
              Fahey, J., Helton, E., Kesteman, M., Madan, A., Young, A.C., Shevchenko, Y.,
              Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
              Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
              Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
              Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smalys, D.E.,
              Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
              Generation and initial analysis of more than 15,000 full-length
              human and mouse cDNA sequences
              Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
              12477932
              2 (bases 1 to 1813)
              Strausberg, R.
              Direct Submission
              Submitted (02-JUN-2003) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
              USA

```

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Leonard I. Zon, M.D.
 DNA Library Preparation: Invitrogen Corp
 DNA Sequencing by: The I.M.A.G.E. Consortium (LILNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: <http://image.lilnl.gov>
 Series: IRAC Plate: 117 Row: e Column: 18
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

Location/Qualifiers

source

```

1..1813
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:63960 IMAGE:6791907"
/tissue_type="Kidney, zebrafish"
/clone_1ib="NCI CGAP_ZK1d1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
1..1813
/gene="zgc:63960"
/note="synonym: MGC63960"
/db_xref="LocusID:393322"
178..786
/gene="zgc:63960"
/codon_start=1
/product="hypothetical protein MGC63960"
/protein_id="AAH53167.1"
/db_xref="GI:31418811"
/db_xref="LocusID:393322"
/translatiion="MDSGGESEELSPGHESFNKSIALOAKKILSMATMAVAN
LITDDISSLLELRYASRYTSKKAHKITIDVYIKALKITILYRNHOPSEBNET
VERFKKMOAATVVSFYEVETPDGILSLLELCRDLHELVHILTRSHGRID
HFNHPADVDFILEYPSBEDYRLNLRKICDGNKLIDBGLT"

```

gene

CDS

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3.36e-44	84.40	60.44%	58.89%	62.37%	5	1813	106	46	28	0	0

ORIGIN

US-10-627-571-2 (1-188) x BC053167 (1-1813)

```

Qy      6  PheAsnSerLysAsnLeuAlaGlnAlaGlnLysLysIleLeuGluYsMetValSer 25
Db      235  TTCAATTCCAAAGATTGGCCCTTCAGGCTCAAAAGAAGATTGTGAGTAAATGGCCACC 294
Qy      26  LysSerIleAlaThrLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 45
Db      295  ATGGCGGTGCGAATCTCTCAACAGACGACACACAGCAGCAGATTCGACGAATCTTAC 354
Qy      46  ArgValIthArggluYrThrGlnAsnLysLysLysLysLysLysLysLysLysLys 65
Db      355  AAGGCCAGTGGAAATTAACCAAGACAGAAAGAAAGCCACAAAGATCATCAAGAGATGTC 414
Qy      66  IleYsThrValIleYsLysLysLysLysLysLysLysLysLysLysLysLysLys 85
Db      415  ATCAAGATGCTGTGAAGATGGCATTTCTTACCGGAACACACAGATTCAGTCTGAGAG 474
Qy      86  LeuAlaLeuMetGluYsPheYsLysLysLysLysLysLysLysLysLysLysLys 105

```

Db	475	ATGAGACAGTGGAGCGCTTC	AAAAAGAAATGAACCAAGCGGC	CATTGACGTGGTGACC	534
Oy	106	PheHISGlnValAspTyrThr	PheAspArgAsnValLeuSer	ArgLeuLeuAsnGluCys	125
Db	535	TTTATGAAAGTGAGACATTC	CAACCAAGGATTCCTTCAAG	GTGCTGGTGGATGT	594
Oy	126	ArgGluMetLeuHISGlnIle	IleGlnArgHISLeuThr	AlaLeuSerHISGlyArgVal	145
Db	595	AGAGACCTTCTCCAGAGCG	TGGTGAGCACTTACCATGCG	CGCTCACACGGCGCGGATT	654
Oy	146	AsnAsnValPheAspHIS	PheSerAspCysGluPhe	LeuAlaAlaLeuTyrArgAsnPro	165
Db	655	GACCACTTTTAAACCATTT	CGCGATGTGGATTTCTTCA	CCGACGCTGAGGCCCATCT	714
Oy	166	GlyAsnPheLeuProHIS	LeuGlnIleLeuCysAsp	GlyTyrLeuAsnMetLeuAsp	185
Db	715	GAAAGCTACAGCTTAACT	CTGAGGAAGATCTGGATGG	AGATAAACAACTCTTACG	774

RESULT	43
LOCUS	CR407586.c
DEFINITION	CR407586 134497 bp DNA linear HTG 26-AUG-2004
ACCESSION	Danio rerio clone CH211-28366, WORKING DRAFT SEQUENCE.
VERSION	CR407586
KEYWORDS	HTG; HTGS PHASE2; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. 1 (bases 1 to 134497) Wood,J. Direct Submission Submitted (25-AUG-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Aug 26, 2004 this sequence version replaced gi:50871872.
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

```

Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish\_help@sanger.ac.uk
----- Project Information -----
Center project name: zc283g6
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 134439 bases at least Q40
Consensus quality: 134480 bases at least Q30
Consensus quality: 134490 bases at least Q20
Insert size: 136497; sum-of-contigs
Insert size: 136131; 6.0% error; agarose-fp
Quality coverage: 10.02x in Q20 bases; sum-of-contigs Quality
coverage: 9.90x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```

```

FEATURES
    source
        location/Qualifiers
            1..134497
                /organism="Danio rerio"
                /mol_type="genomic DNA"
                /db_xref="taxon:7955"
                /clone="CH211-263G6"
                /clone_lib="CHOR-211"
            1..134497
                /note="assembly_fragment:001511
                clone_end:17
                vector_side:right"

```

Alignment Scores:	
Pred. No.:	2.85e-42
Score:	600.00
Percent Similarity:	84.44%
Percent Mismatch:	15.56%
Best Local Similarity:	58.89%
Query Match:	62.37%
DB:	2
Gaps:	0
Length:	14497
Matches:	106
Conservative:	48
Mismatches:	26
Indels:	0
Gaps:	0

US-10-627-571-2 (1-188) x CR407586 (1-134497)

OY	6	PheAsnSerIysAsnIleuAlaValGlnIaGlnIysLysIleuGluIlysmMetAlaSer	25
Db	26322	TTCAATTCCAAAAGTTGGCCCTTCAGGCTCAAAAGAAATTTGGCAAAATGGCCACC	26263
OY	26	LysSerIleAlaThrThrIleuIleAspSerThrSerGluValIleuAspGluLeuTyr	45
Db	26262	ATGGCCCTGGCGAAACCTCCTTAACAGACGACACACGAGCGAATTGTGACCAACTTAC	26203
OY	46	ArgValThrArgIuTyrThrGlnAsnLysLysGlnAlaGluLysIleLysAsnLeu	65
Db	26202	AAGCCAGATCGAGAAATACACCAAGACCAAGAAAGAAAGCCACACAGATCTCAAAAGATTC	26143
OY	66	IleLysThrValIleLysIleuAlaIleLeuTyrArgAsnAsnIleAsnGlnIleAspGlu	85
Db	26142	ATCAAGATCGCCTCGAAGATTCGATCTTCTACCGGAACCAACAGATTCCTGTGGAG	26083

86 LeuAlaLeuMetGlyLysPheLysLysValHisGlnLeuAlaMetThrValValSer 105

Db 26082 ATGAGACCGTGGAGCGCTTCAAAAAGAGATGAACCGCGGCCATGACAGTGGTGAGC 26023

QY 106 phenIsglnvalasptrythrphenaspargasnvalleuserargleuleuasnglucys 125

Db 26022 TTTATGAGTGGAGTACACATTGACCGAGCATCCTTCAGAGCTGCTGTGGAATGT 25963

QY 126 ArgGlutMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal 145

Dp 25962 AGAGACCTTCTCCACGAGCTGTTGAGCACCACTTGACCATGCGCTCCCACGGCGAATT 25903

146 AsnAsnVal¹pheAsnHis²pheSerAspCysGly³pheLeuAla⁴AlaLeuTyrAsnPro⁵phe 165

dh 25902 GACCAAGTTTTCACACCAATTTTCGCCGATGTCGATTTTCTTGACCGAGCTGTACCGGCCCATCTT 25843

166 C]vAenBhai.vBBrCh'i.eI.aW[C]nI.v.eI.aW.CvAenC]vT]eAeN.I.v.eM.eT.I.euB.eeC]" 185

35783 35843

SECRET

BK927313/c
NYO05333
100507 h-
NYO
11-1-1994
JAN 06 09M 2004

DEFINITION Zebrafish DNA sequence from clone CH211-12A1 in linkage group 18,

ACCESSION BX927313

KEYWORDS	HTG.
VERIFICATION	01.07000000
DATE/2010.11	01.07000000

SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

REFERENCE
1 (bases 1 to 189797)
Cypripiniiformes; Cypripinidae; Danio.

AUTHORS	Mashreghni-Monammadi, M.
TITLE	Direct Submission

JOURNAL
Submitted (08-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

COMMENT
On Oct 9 2004 this sentence version replaced at:52137684
zfsh-help@sanger.ac.uk clone requests: clonerequest@sanger.ac.uk

----- Genome Center
Center: Wellcome Trust Sanger Institute

Center code: SC
Web site: <http://www.sanger.ac.uk>

Contact: zf1sh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

126 ArgGluMetLeuHisglnIleIleIleGlnArgHisLeuThrAlaValSerHisglnIleArgVal 145
DD 10/3/4 ITTIAIGAGIGGAGIACACAITCGACCGAGGCAITCTTTGAGAGCTGCTGTGAGTAIT 167

```
repeat_region  
complement(219/.2502)  
/rpt_family="AluSx"  
repeat_region  
3589.3609
```

repeat_region	4266..4572 /rpt_family="AluJo"
repeat_region	4692..4784 /rpt_family="MIR"
repeat_region	4793..4874 /rpt_family="MIR"
repeat_region	compLement(4888..5232) /rpt_family="AluJo"
repeat_region	compLement(5239..5545) /rpt_family="AlusX"
misc_feature	5741..5848 /note="Predicted exon, program: gffutils_human_1.3, frame: 1, quality: good, score: 61.000"
repeat_region	compLement(6054..6181) /rpt_family="FLAM_C"
repeat_region	compLement(6188..6262) /rpt_family="MERSA"
repeat_region	6677..6826 /rpt_family="GC_rich"
repeat_region	compLement(6973..7032) /rpt_family="(CAGA)n"
repeat_region	7764..7895 /rpt_family="MIR"
repeat_region	7919..8022 /rpt_family="LINE2"
repeat_region	8827..8868 /rpt_family="MIR"
repeat_region	9104..9184 /rpt_family="MIR"
repeat_region	9219..9441 /rpt_family="AluJb"
repeat_region	9444..9743 /rpt_family="AlusX"
repeat_region	compLement(9773..10218) /rpt_family="LINE2"
repeat_region	10246..10538 /rpt_family="AlusX"
repeat_region	10911..11179 /rpt_family="AlusX"
repeat_region	11184..11484 /rpt_family="AlusP"
repeat_region	11503..11807 /rpt_family="AluJo"
repeat_region	compLement(11856..12145) /rpt_family="AlusX"
repeat_region	12511..12840 /rpt_family="AlusX"
repeat_region	12868..12968 /rpt_family="AluJ/FLAM"
repeat_region	13135..13228 /rpt_family="LINE2"
repeat_region	compLement(13408..13671) /rpt_family="AlusQ"
repeat_region	compLement(13674..13880) /rpt_family="LMB7"
repeat_region	compLement(13881..14172) /rpt_family="AluY"
repeat_region	compLement(14175..14555) /rpt_family="LMB7"
repeat_region	compLement(14560..14854) /rpt_family="AlusX"
repeat_region	compLement(14865..15172) /rpt_family="AluJo"
repeat_region	15228..15275 /rpt_family="LINE2"
repeat_region	compLement(16172..16474) /rpt_family="AlusQ"
repeat_region	compLement(16664..16716) /rpt_family="MIR"
repeat_region	16853..16917 /rpt_family="MIR"
repeat_region	18083..18246 /rpt_family="AluY"

```

repeat_region      complement (18247, .18276)
                    /rpt_family="AT_rich"
repeat_region      18551, .18586
                    /rpt_family="AT_rich"
repeat_region      complement (18596, .18711)
                    /rpt_family="FLAM_c"
repeat_region      complement (18725, .18931)
                    /rpt_family="Alusq"
misc_feature        19115, .19675
                    /note="predicted exon, program: grail2exons human_1.3,
                    frame: 1, quality: excellent, score: 92.000"
repeat_region      20087, .20262
                    /rpt_family="MLT1C"
repeat_region      20269, .20558
                    /rpt_family="Aluv"
repeat_region      20575, .20711
                    /rpt_family="Aluv"
repeat_region      20720, .21034
                    /rpt_family="Alusc"
repeat_region      21049, .21201
                    /rpt_family="Aluv"
repeat_region      21225, .21398
                    /rpt_family="MLT1C"
misc_feature        complement (21329, .21549)
                    /note="BLASTN similarity to Z42385 (1, .220); match: 0.97,
                    score: 9.7e-83; database searched: est; H. sapiens partial
                    cDNA sequence"
misc_feature        21389, .21554
                    /note="predicted exon, program: grail2exons human_1.3,
                    frame: 1, quality: excellent, score: 77.000"
repeat_region      complement (21587, .21852)
                    /rpt_family="Alub"
repeat_region      22809, .22929
                    /rpt_family="Alub"
repeat_region      22943, .23243
                    /rpt_family="Aluv"
repeat_region      23246, .23404
                    /rpt_family="Alusx"
repeat_region      complement (23833, .24137)
                    /rpt_family="Aluao"
repeat_region      24185, .24484
                    /rpt_family="Alub"
misc_feature        complement (24790, .25329)
                    /note="UDS similarity to overlapping ESTs:
                    (25329, .25201) AA477269 zu43c12.s1 Soares ovary tumor
                    NbhOT Homo sapiens cDNA clone 740758.3'; (401, .273) 100%
                    identity.~(25329, .25201) A477268 zu43c12.r1 Soares ovary
                    tumor NbhOT Homo sapiens cDNA clone 740758.5'; (221, .349)
                    100% identity.~(25061, .24842) A477268 zu43c12.r1 Soares
                    ovary tumor NbhOT Homo sapiens cDNA clone 740758.5';
                    (350, .570); 98% identity.~(24792, .25272) N13339 yw22908,
                    Homo sapiens cDNA clone 258782.3'. Score: 925 Identity:
                    476/480 (99%) ~(.25329, .25028) AA579149 nF28a04.s1
                    NCI CGAP P71 Homo sapiens cDNA clone IMAGE:915054;
                    (134, .434); 99% identity.~(24790, .25061) AA477269
                    zu43c12.s1 Soares ovary tumor NbhOT Homo sapiens cDNA
                    clone 740758.3'; (272, .1); 100% identity.~Additional EST
                    matches:
                    AA581955, AA467935, AI038745, AI041764, T24716"
                    complement (561n1(25250, .25329, 27176, .27248, 27914, .27995,
                    32121, .32185))
                    /note="Hypothetical partial human protein"
                    /codon_start=1
                    /evidence=not experimental
                    /product="R33729.1"
                    /protein_id="AAC77824.1"
                    /db_xref="GI:3355455"

Alignment Scores:
Pred. No.:          2  77e-42      Length:      32360
Score:              593.00      Matches:         106
Percent Similarity: 79.14%      Conservative: 42
Best Local Similarity: 56.68%      Mismatches:    39

```

```
Query Match: 61.64% Indels: 0
DB: 9 Gaps: 0
```

US-10-627-571-2 (1-188) x AC005339 (1-32360)

QY	2	AlaThrAspValPheasnSerIysAsnLeuAlaValGlnIleGlnIleuGly	21
Db	19112	GCATGACACCTTCAGACCAACAGCCTGGCTTCAGAGCGCAAGAAAGCTCTGAGT	19177
QY	22	LysMetValSerIysSerIleAlaThrThreIleAspPheSerSerGluValIleu	41
Db	19172	AAAGATGGCGTCCAAAGCAGCGTGGTGGCGCTGGTGGATACACCAAGACTGAGTGGT	19233
QY	42	AspGluLeuTyTArgValThrArgGluTyThrGlnAsnIysGluAlaGluIysIys	61
Db	19232	GATAGAGCTGATCCCGCGCACCGAGGAGTTGACGCGCAGCGCAAGGAGCCGCAAGAGT	19291
QY	62	IleIysAsnIleuIleTyThrValIleIysLeuAlaIleIleuTyArgAsnAsnIlePhe	81
Db	19292	CTCAAGAACCTGGTCAAGGTGGCCCTTAAGCTGGAGACTGGCTGGCTGGAGCCAGCTGG	19351
QY	82	AsnGlnAspGluLeuAlaIleMetGluTyPheIleValIleGlnLeuAlaMet	101
Db	19352	GGCGGTAGAGAGCTGGCGGTGGTGGCGGCTTCGCGCACCGGCGCGGCTCGTGGCAG	19411
QY	102	ThrValValSerPheHISGlnValAspTyThrPheAspArgAsnValIleuSerArgIleu	121
Db	19412	ACGGCCGTCACCTTCACACAGTGTGACTTCACTTCACACGGCGCGTGGCGCGCCGG	19471
QY	122	LeuAsnGluCysArgGluMetIleuHISGlnIleIleGlnIleArgIleuThrAlaIysSer	141
Db	19472	CTGCTCGAGTCCCGGACCTGGCTGCACACAGCCGCTGGTCCCACTGACCGCAAGTCC	19531
QY	142	HISGluTyArgValAsnAsnValIleAspArgIlePheSerAspCysGluPheLeuAlaIleu	161
Db	19532	CACGGCCGCACTAACCAACAGTGTGGCGGACCTAGCCGACCTTCCTGGCTGGCGCTC	19591
QY	162	TyrAsnProPheGluAsnPheIysProHISLeuGlnIleuIleuCysAspGluIleAsnIle	181
Db	19592	TACGGCCCGCGAGCCCTACCGCTCCACCTGGCGAGATGTGAGAGGCGCTGGGCGG	19651
QY	182	MetIleuAspGluGluAsnIle	188
Db	19652	ATGCTGACAGAGGCAAGCTTC	19672

Search completed: July 28, 2005, 20:50:53
Job time : 4523 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 18:13:05 ; Search time 175 Seconds
(without alignments)
1757.825 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962
Sequence: 1 MATDVNSKNINVAQAQKIL.....KPHLQKLCDCGINKMLDEBN1 188

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgnt2_1/USPTO.spool.h/US10627571/runat.27072005.154721.27515/app.query.fasta_1.327
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPTOL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humano.cdi
-LIST=100 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10627571@CGN1.1.69@runat.27072005.154721.27515 -MCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgnt2_6/prodata/1/ina/5A.COMB.seq.*
2: /cgnt2_6/prodata/1/ina/5B.COMB.seq.*
3: /cgnt2_6/prodata/1/ina/6A.COMB.seq.*
4: /cgnt2_6/prodata/1/ina/6B.COMB.seq.*
5: /cgnt2_6/prodata/1/ina/PCRTUS.COMB.seq.*
6: /cgnt2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	527	54.8	1175	4 US-09-799-451-537	Sequence 537, App
2	250	26.0	340	4 US-09-513-999C-14427	Sequence 14427, A
3	166	17.3	252	4 US-09-513-999C-14434	Sequence 14434, A
4	104	10.8	1347	4 US-09-248-796A-6176	Sequence 6176, Ap
5	96.5	10.0	2229	4 US-09-248-796A-4950	Sequence 4950, Ap
6	92.5	9.6	2253	4 US-09-949-016-3908	Sequence 3908, Ap
7	90.5	9.4	5923	4 US-09-976-594-907	Sequence 907, App
8	90	9.4	1887	4 US-09-601-198-39	Sequence 39, App
9	90	9.4	1664976	4 US-08-916-421B-1	Sequence 1, Appli
10	89.5	9.3	1142	4 US-09-175-684A-8	Sequence 8, Appli
11	89.5	9.3	1975	4 US-09-949-016-975	Sequence 975, App

13	88.5	9.2	1700	4 US-09-016-434-1081	Sequence 1081, Ap
14	88.5	9.2	1700	4 US-09-023-655-898	Sequence 898, App
15	88.5	9.2	1865	4 US-09-270-767-14175	Sequence 14175, A
16	88.5	9.2	3204	4 US-09-710-279-3485	Sequence 3485, Ap
17	88.5	9.2	3627	4 US-09-710-279-4311	Sequence 4311, Ap
18	88.5	9.2	3760	4 US-09-710-279-4029	Sequence 4029, Ap
19	88	9.1	1638	4 US-09-616-289-46	Sequence 46, Appl
20	88	9.1	4206	4 US-09-750-590A-3	Sequence 3, Appli
21	88	9.1	4637	3 US-08-979-608A-17	Sequence 17, Appl
22	88	9.1	4697	4 US-09-517-849-17	Sequence 17, Appl
23	88	9.1	4697	4 US-09-516-289-17	Sequence 17, Appl
24	88	9.1	4730	4 US-09-750-590A-1	Sequence 1, Appli
25	87.5	9.1	3172	4 US-09-976-594-940	Sequence 940, App
26	87.5	9.1	3349	3 US-09-336-447A-2	Sequence 2, Appli
27	87.5	9.1	3349	3 US-09-352-267B-2	Sequence 2, Appli
28	87.5	9.1	4722	3 US-08-979-608A-14	Sequence 14, Appl
29	87.5	9.1	4722	4 US-09-517-849-14	Sequence 14, Appl
30	87.5	9.1	4722	4 US-09-516-289-14	Sequence 14, Appl
31	87	9.0	4223	3 US-09-541-782-5	Sequence 5, Appli
32	87	9.0	4223	3 US-09-723-820-5	Sequence 5, Appli
33	87	9.0	4223	4 US-10-270-085-5	Sequence 5, Appli
34	86.5	9.0	2469	4 US-09-248-796A-4746	Sequence 4746, Ap
35	86	8.9	1299	4 US-09-248-796A-5857	Sequence 5857, Ap
36	86	8.9	2651	1 US-08-462-949-1	Sequence 1, Appli
37	86	8.9	2651	1 US-08-023-764B-1	Sequence 1, Appli
38	85.5	8.9	618	4 US-09-134-000C-1489	Sequence 1489, Ap
39	85.5	8.9	1854	5 PCT-US94-01101-1	Sequence 1, Appli
40	85.5	8.9	8590	4 US-09-949-016-5562	Sequence 5562, Ap
41	85.5	8.9	10300	4 US-09-949-016-636	Sequence 636, App
42	85	8.8	2637	3 US-09-134-001C-1541	Sequence 1541, Ap
43	84.5	8.8	571	3 US-09-404-879A-84	Sequence 84, Appl
44	84.5	8.8	571	4 US-09-338-933-84	Sequence 84, Appl
45	84.5	8.8	571	4 US-09-215-681-84	Sequence 84, Appl
46	84.5	8.8	571	4 US-09-216-003A-84	Sequence 84, Appl
47	84.5	8.8	571	4 US-09-667-857-84	Sequence 84, Appl
48	84.5	8.8	3025	4 US-09-917-725A-19	Sequence 19, Appl
49	84	8.7	906	4 US-09-248-796A-4219	Sequence 4219, Ap
50	84	8.7	1485	4 US-09-248-796A-53	Sequence 53, Appl
51	84	8.7	58407	4 US-08-916-421B-2	Sequence 2, Appli
52	84	8.7	58407	4 US-09-692-470-2	Sequence 2, Appli
53	84	8.7	640681	4 US-09-790-988-1	Sequence 1, Appli
54	83.5	8.7	734	4 US-08-956-171E-1051	Sequence 1051, Ap
55	83.5	8.7	734	4 US-08-781-986A-1051	Sequence 1051, Ap
56	83.5	8.7	833	4 US-09-134-000C-1958	Sequence 1958, Ap
57	83.5	8.7	1065	4 US-09-175-684A-1	Sequence 1, Appli
58	83.5	8.7	1088	4 US-09-175-684A-2	Sequence 2, Appli
59	83.5	8.7	1575	4 US-09-543-681A-615	Sequence 615, App
60	83.5	8.7	1785	4 US-09-117-415B-1	Sequence 1, Appli
61	83.5	8.7	1862	4 US-09-117-415B-21	Sequence 21, Appl
62	83.5	8.7	1896	4 US-09-117-415B-17	Sequence 17, Appl
63	83.5	8.7	1896	4 US-09-117-415B-19	Sequence 19, Appl
64	83.5	8.7	1950	4 US-09-117-415B-15	Sequence 15, Appl
65	83.5	8.7	2949	4 US-09-023-905A-8	Sequence 8, Appli
66	83.5	8.7	3446	4 US-09-392-714-13	Sequence 13, Appl
67	83.5	8.7	3636	3 US-09-134-001C-1983	Sequence 1983, Ap
68	83.5	8.7	3954	4 US-09-023-905A-6	Sequence 6, Appli
69	83	8.6	2553	4 US-09-248-796A-2007	Sequence 2007, Ap
70	83	8.6	2553	4 US-09-269-858A-7	Sequence 7, Appli
71	83	8.6	14853	2 US-08-793-824-1	Sequence 1, Appli
72	83	8.6	14736	3 US-08-961-527-171	Sequence 171, App
73	83	8.6	1664976	4 US-08-916-421B-1	Sequence 1, Appli
74	83	8.6	1664976	4 US-09-692-570-1	Sequence 1, Appli
75	82.5	8.6	5467	4 US-09-976-594-735	Sequence 735, App
76	82.5	8.6	5967	4 US-09-949-016-1162	Sequence 1162, Ap
77	82.5	8.6	6640	4 US-09-949-016-797	Sequence 797, App
78	82.5	8.6	30549	3 US-09-134-001C-322	Sequence 322, App
79	82	8.5	1026	4 US-09-248-796A-12786	Sequence 12786, A
80	82	8.5	1522	4 US-09-620-312D-56	Sequence 56, Appl
81	82	8.5	1159	4 US-09-248-796A-3452	Sequence 3452, Ap
82	82	8.5	8257	4 US-09-595-684B-30	Sequence 30, Appl
83	82	8.5	8503	4 US-09-620-312D-130	Sequence 130, Appl
84	81.5	8.5	906	3 US-09-134-001C-78	Sequence 78, Appl
85	81.5	8.5	968	4 US-08-956-171E-553	Sequence 553, App


```

Alignment Scores:
Pred. No.: 9.39e-24 Length: 340
Score: 250.00 Matches: 52
Percent Similarity: 96.36% Conservative: 1
Best Local Similarity: 94.55% Mismatches: 2
Query Match: 25.99% Indels: 0
DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x US-09-513-999C-14427 (1-340)

QY 2 AlaThrAspValAlpheAnSerLySAsnLeuAlaValGlnAlaGlnLySlySileLeuGly 21
Db 174 GCACGAAATGTCCTTTATTCCTCAAAAACCTGCGCGKTCAGGCACAAAARAAGATCTTG 23
QY 22 LysMetValSerLySserLileAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
Db 234 AAAATGGTGTCCAAAATCCATCGGCACACCTTAATGACACACAAAGTAAGTGAGTGCTG 293
QY 42 AspGluLeuThrArgValThrArgGluThrGlnAsnLySlyS 56
Db 294 GATGACCTCTACAGAGTGACCGAGGAGTACACCCAAAACAAAGAG 338

RESULT 3
US-09-513-999C-14434
; Sequence 14434, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14434
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 32
; OTHER INFORMATION: Y=c or t
US-09-513-999C-14434

Alignment Scores:
Pred. No.: 6.42e-13 Length: 252
Score: 166.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.26% Indels: 0
DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x US-09-513-999C-14434 (1-252)

QY 1 MetAlaThrAspValAlpheAnSerLySAsnLeuAlaValGlnAlaGlnLySlySileLeu 20
Db 148 AAGGGCCACGAGATGCTTTAATTCCAAAAACCTGCGCGTTCAGGCACAAAAGAAGATCTTG 207
QY 21 GlyLysMetValSerLySserLileAlaThrThrLeuIleAspAsp 35
Db 208 GGTAAATGTGTCCAAATCCATCGGCACACCACTTAATAGACGAC 252

RESULT 4
US-09-248-796A-6176
; Sequence 6176, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

```

[illegible]

```
/ Sequence 4950, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 4950
/ LENGTH: 2229
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-4950

Alignment Scores:
Pred. No.: 0.0228 Length: 2229
Score: 96.50 Matches: 44
Percent Similarity: 33.98% Conservative: 26
Best Local Similarity: 21.36% Mismatches: 55
Query Match: 10.03% Indels: 81
DB: 4 Gaps: 6

US-10-627-571-2 (1-188) x US-09-248-796A-4950 (1-2229)
QY 3 ThrAspValIphesAsnSerIysAsnLeuAlaValGlnAlaGlnLys----- 18
DB 1252 ACAAAATCATCATATGTAAGATGAAGATTCACACACACAGAAAGATGAGACAC 1311
QY 19 -----IleLeuGlyLysMetValSerIysSer 27
DB 1312 ATGAAATGAGACTGAAAAAGAGAGTTGATGATTAATGAAAAAATTACAGAAAAAGAT 1371
QY 28 IleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuIleVal 47
DB 1372 GAGGAATGTCAGCTGCTAAACAGAAATAGATGATGATGAGATTCG 1431
QY 48 ThrArg-----Glu 50
DB 1432 ACCAAGAACTGTAAGTCACTGATATATCATTAATGAAGATTAGAACAAAAATTAAGTCA 1491
QY 51 TTTTGTGlnAsnLysLysGluAlaGluLysLysIleLysAsnLeuIleLysThrValIle 70
DB 1492 TACACTCAGAAAGCCACTTCACAAACCAAAAGAAATGCTGATTTAAACAAAGACGTTTCG 1551
QY 71 LysIleu-----AlaIleLeuTyArgAsn----- 78
DB 1552 AATTTCGCTTCGAGCTAGGTGATGCTATGATGATGATGATGATGATGATGATGATGATGAT 1611
QY 79 -----AsnGlnPheAsnGlnAspGluLeuAla 87
DB 1612 TCTTCAAGAGAGCCGAGTTTACTAAGAACGCAATCATTTTACATGATTTGAAGCT 1671
QY 88 LeuMetGlnLysPhe----- 92
DB 1672 TTGCATTTGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1731
QY 93 -----LysLysLysValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyr 111
DB 1732 GCAAAACCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1773
QY 112 ThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGln 131
DB 1774 -----AATGTGATATCATCTTAATGAATGATTTAAGAAACATTATTCAGT 1821
QY 132 IleIleGlnArgHisLeu 137
DB 1822 GTTTTGAATATTTTTCG 1839
```

```
RESULT 6
US-09-949-016-3908
/ Sequence 3908, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR FILING DATE: 2000-09-08
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3908
/ LENGTH: 2253
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-3908

Alignment Scores:
Pred. No.: 0.0777 Length: 2253
Score: 92.50 Matches: 52
Percent Similarity: 40.00% Conservative: 36
Best Local Similarity: 23.64% Mismatches: 73
Query Match: 9.62% Indels: 59
DB: 4 Gaps: 12

US-10-627-571-2 (1-188) x US-09-949-016-3908 (1-2253)
QY 8 SerLysAsnLeuAlaValGlnAlaGlnLys-----Lys 18
DB 244 GCCCGAAGCTTCAAGGTTTACCTGCAGAGGAGCTTAATTCACCTGATGATCTGTAATA 303
QY 19 IleLeu-----GlyLysMetValSerLysSerIleAlaThrThrLeuIle 33
DB 304 ATCTCAGCTCTTCCAGATGGAAGAAATGTTGAAATTTGATGATGCTGTACTGAGATT 363
QY 34 AspAspThrSerSerGluValLeuAspGluLeuIleValThrValThrArgGluLysThrArg 53
DB 364 CTAAATTAATGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 423
QY 54 AsnLysLysGluAlaGluLysLysIleLysAsnLeuIleLysThrValIleLysLeuAla 73
DB 424 GTAGCAAAAGCGGCTATCAATCCCTGTCAGAAATATCATCAATCAATCAATCAATCAATCA 483
QY 74 IleLeuTyArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGlnLys----- 91
DB 484 GCTTATTTGAAGCAATCTGCTG-----GATGATTTGAAAAGGTAAATGAACAAAT 537
QY 92 -----PheLysLysLysValHisGlnLeuAlaMetThrVal-----ValSerPheHis 107
DB 538 GACATTTGTCATACAGGCTGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 597
QY 108 GlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGlu----- 124
DB 598 TCTTTAACTACTGTACACCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 657
QY 125 -----CysArgGluMetLeuHisGlnIleIleGlnArgHis 136
DB 658 GAGGATGTGTGTGTCAGAGCCAGCTGTATAGAAATGTGCATCACTGACATATATCTACTAT 717
QY 137 LeuThrAlaLysSerHisGlyArg-----ValAsnAsnValPheAspHisPhe 152
DB 718 -----CATGGCGAGCAATATCTTGTCTCAAGAGAGATTAATTCACCAAT 762
QY 153 Ser-----AspCysGlnPheLeuAlaIleLeuTyArgAsnPro----- 164
```


Db 1026 AAAAAATCA-----TTAGATATTGAACAATT 1000
Qy 159 A1a1a1eUtyrAsnProhPheGlyAsnPhelyProhiIsuenglInysleuCySaBpGly 178
Db 999 AATCTTTAACTNATATAATATTATTAAT-----GATTACAAATAATTATTAACAAGT 949
Qy 179 I1eAsnlyAmE---LeuSpGluGluAAsn 187
Db 948 TTAATATCAATTCATTAAGTACTGTAATAAT 919
RESULT 9
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98256)..(98256)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)

OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g

```
/ NAME/KEY: misc_feature
/ LOCATION: (1084830)..(1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1096846)..(1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1119881)..(1119881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1130881)..(1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1345473)..(1345473)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1349491)..(1349491)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1470091)..(1470091)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1569020)..(1569020)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1602912)..(1602912)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1603734)..(1603734)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1637998)..(1637998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (1664854)..(1664854)
/ OTHER INFORMATION: n equals a, t, c, or g
/ OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Alignment Scores:
Pred. No.: 3.42e+03 Length: 1664976
Score: 90.00 Matches: 43
Percent Similarity: 43.85% Conservative: 39
Best Local Similarity: 22.99% Mismatches: 64
Query Match: 9.36% Indels: 41
DB: 4 Gaps: 9

US-10-627-571-2 (1-188) x US-08-916-421B-1 (1-1664976)
QY 9 lyeAsnLeuAlaVal-----GlnAlaGlnLysLys 18
DB 275458 AAAAAGCTTGCAATATAAATAAAGATTGGAGATTAATTAAGCTAATTAATAAAC 275399
QY 19 lIeLeuGlyLysMetValSerLysSerIleAlaThrThrLeuIleAspThrSerSer 38
DB 275398 TTGCTTAATTAAGATGAGATTATTAAGTGTCTTAAGTAAAGAAAGATTTCAGATATGAAGT 275339
QY 39 GluVal-----LeuAspGlu-----LeuTyrArgValThrArgGlu----- 50
DB 275338 CAATATCAAAAAGCTTTAGATGAATAATATATATATAAGAAATAATTGATCATTAAT 275279
QY 51 ---TyrThrGlnAsnLysLysGluAlaGlnLysLysIleLysAsnLeuIleLysThrVal 69
DB 275278 AATATATATCGAAACATTAAGAAAGAAAGATGATTAACCTAAAGATTAAGAGTTAAGAGTTA 275219
QY 70 lIeLysLeuAlaIleLeuTyrTyrGlnAsnGln-----PheAsn 82
DB 275218 AGTGATATTCGAAATAATATCATGGAAGAAAGAGAGTATTAAGAGCTTTTAAC 275159
```

```
QY 83 GlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThr 102
DB 275158 GAAGATGAATAAATTTATTCGATTAAATTAATAAATAATGACATATTAACCAAAAAA 275099
QY 103 ValValSerPheHisGlnValAlaSerTyrThrPheAspArgAsnValLeuSerArgLeu 122
DB 275098 ATTGTTGAATTAATCTCGAATG-----AGCAAAACCAAGGTTTCGAGAAATTAATC 275051
QY 123 AsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHis 142
DB 275050 TCCGAATTAAGAGAT-----AGAAATAATATAGGAAA-----GAAAAATA 275009
QY 143 GlyArgValAsnAsnValPheAspHisPheSerAspCysGlu-PheLeuAlaLeuTyr 162
DB 275008 GGACGAATCAATTAATTA-----ACCTTAACAGAAAGAGTAAGAAAGTTATTA 274961
QY 162 rAsnProPheGlyAsnPhe 168
DB 274960 TAACCAATTAATAATTTT 274942

RESULT 10
US-09-692-570-1/c
/ Sequence 1, Application US/09692570
/ Patent No. 6797466
/ GENERAL INFORMATION:
/ APPLICANT: Buit et al.
/ TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-
/ Patent No. 6797466
/ TITLE OF INVENTION: jannaschi
/ FILE REFERENCE: PB275C1
/ CURRENT APPLICATION NUMBER: US/09/692,570
/ PRIOR FILING DATE: 2003-01-14
/ PRIOR APPLICATION NUMBER: US 60/024,428
/ PRIOR FILING DATE: 1996-08-22
/ PRIOR APPLICATION NUMBER: US 08/916,421
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1664976
/ TYPE: DNA
/ ORGANISM: Methanococcus jannaschi
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28222)..(28222)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28257)..(28258)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84773)..(84773)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84812)..(84812)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98120)..(98120)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98159)..(98159)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
```

LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)

OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

Alignment Scores:
Pred. No.: 3,42e+03 Length: 1664976
Score: 90.00 Matches: 43
Percent Similarity: 43.85% Conservative: 39
Best Local Similarity: 22.99% Mismatches: 64
Query Match: 9.36% Indels: 41
DB: 4 Gaps: 9

US-10-627-571-2 (1-188) x US-09-692-570-1 (1-1664976)

QY 9 LysAsnLeuAlaVal-----GlnAlaGlnLysLys 18
Db 275458 AAAAAGTTACCAATAAATAAAGATTGGAGATAATTAAGTAAAGGCTAATATAAAC 275399

QY 19 IleLeuGlySerMetValSerLysSerIleAlaThrThrIleuIleAspAspThrSerSer 38
Db 275398 TTGCTAATATGAAGATGAGATTATTAGTCTTAATGAAGATTTCAGAAATGAAAGT 275339

QY 39 GluVal-----LeuAspGlu-----LeuTyArgValThrArgGlu----- 50
Db 275338 CAAATACAAAACCTTTAGATGAAATAATATATATATAGAAAATAATGAATTCATTAAAT 275279

QY 51 ---TyThrGlnAsnLysLysGluAlaGlnLysIleLysAsnLeuIleLysThrVal 69
Db 275278 AATATATGCAACATTAAAGAAAGAAATGATAGCTAAAGATAAGATTAGAGACTTA 275219

QY 70 IleLeuLeuAlaIleLeuTyArgAsnAsnGln-----PheAsn 82
Db 275218 AGTGAATTTCCAAAAAATATCATGGAAGAAAAGAGATATTATGAGCTTTTAACT 275159

QY 83 GlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThr 102
Db 275158 GAAAGATGAAAAATTATATATCGATTATATATAAACAATGACGACTATTACTCAAAAAGAA 275099

QY 103 ValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeu 122
Db 275098 ATTGTTGAATTTACTGSAAG-----AGCAAAACCAAGGTTTGAGAAATTATC 275051

QY 123 AsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHis 142
Db 275050 TCCGAATTAGAAAGAT-----AGAAAAATATATAGGAAA-----GAAAAATA 275009

QY 143 GlnArgValAsnAsnValPheAspHisPheSerAspCysGlu--PheLeuAlaIleLeuTy 162
Db 275008 GGAAGCAATCAATAAATTA-----ACCTTACAGAGAGAGATAAAGATTATTA 274961

QY 162 AsnProPheGlyAsnPhe 168
Db 274960 TAAACCATTTAAAAATTTT 274942

RESULT 11
US-09-175-684A-8
; Sequence 8, Application US/09175684A
; Patent No. 6593463
; GENERAL INFORMATION:
; APPLICANT: Chen, Li How
; APPLICANT: Meade, Harry M.
; TITLE OF INVENTION: NOVEL MODIFIED MSP-1 NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN
; FILE REFERENCE: 10275-133001
; CURRENT APPLICATION NUMBER: US/09/175,684A
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 60/085,649
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/062,592
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 19

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1142)
; US-09-175-684A-8

Alignment Scores:
Pred. No.: 0.0693 Length: 1142
Score: 89.50 Matches: 50
Percent Similarity: 38.02% Conservative: 23
Best Local Similarity: 26.04% Mismatches: 84
Query Match: 9.30% Indels: 35
DB: 4 Gaps: 7

US-10-627-571-2 (1-188) x US-09-175-684A-8 (1-1142)

QY 7 AsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGlyLysMetValSerLys 26
Db 487 AACGAGAGTACCTGCGCTTCTGAAACAATGAGACCTGTACAAAGACCTCAAGAT 546

QY 27 SerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluLeuTyArg 46
Db 547 AACATTGATCTGTTCTGTATC---CACCTGAGCGCAAGGCTCTG----- 588

QY 47 ValThrArgGluTyThrGlnAsnLysLysGluAlaGlnLysLysIleLys-----Asn 64
Db 589 -----CAGTACACATATGAGAAAGCAACGTGAGAGGTCAAGATCAAGAGCTGAT 639

QY 65 LeuIleLysThrVal---IleLysLeuAlaIleLeuTyArgAsnAsnGlnPhe----- 81
Db 640 TACCTGAGACCATCCAGATAGCTGCGGATTTCAGAGAAACAACATTGCTCGGA 639

QY 82 -----AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLys 95
Db 700 ATGCCGATCTGAGCAACGATTACAAACCAACACACCTGCTGACCAAGTTC----- 750

QY 96 ValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyThrPheAspArg 115
Db 751 -----CTGAGACCGGAAATGCTCTTCAAAAACCTGGCCCAAG 786

QY 116 AsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArg 135
Db 787 ACCGTCTTGAACAACCTGCTGATGGAACCTGCAAGATGCTGCAATCAGCCAGAC 846

QY 136 HisLeuThrAlaLysSerHisGlnArgValAsnAsnValPheAspHisPheSerAspCys 155
Db 847 CAGTGTGTAAGAGAGAGTGTCCCAAGACAGCGATGCTTACAGACCTGATGAGAG 906

QY 156 GluPheLeuAlaIleLeuTyArgAsnProPheGlyAsnPheLysProHisLeuGlnLysLeu 175
Db 907 GAGAGAGTCAAGTGCCTGCTG-----AATCTAAGAGAGAGATAAAGTGT 954

QY 176 CysAspGlyIleAsnLysMetLeuAspGluGluAsn 187
Db 955 GTGGAAGAACCCCAATCTACTTGTATACGAGAACAT 990

RESULT 12
US-09-949-016-975
; Sequence 975, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```

; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FaSeq for Windows Version 4.0
 ; SEQ ID NO 975
 ; LENGTH: 1975
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1975)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-949-016-975

Alignment Scores:
 Pred. No.: 0.184 Length: 1975
 Score: 89.00 Matches: 42
 Percent Similarity: 37.50% Conservative: 30
 Best Local Similarity: 21.88% Mismatches: 70
 Query Match: 9.25% Indels: 50
 DB: 4 Gaps: 6

US-10-627-571-2 (1-188) x US-09-949-016-975 (1-1975)

QY 4 AspValPheAsnSerIleuAlaGlnAlaGlnLysIleLeuGlyLysMet 23
 DB 1169 GATGCCGAGAGAGATTATGATGATGGCGGTGGAAGCGAGCGGAGCCGAGCCGAGCCGCTC 1228
 QY 24 ValSerLysSerIleAlaThrThrLeuIleAspAspThrSer-----SerGluVal 40
 DB 1229 CTCAGAGAGCGCGGTGCGAGCCGAGCGGCTGTCGACAGTCAGATGATGATACCCAGATC 1288
 QY 41 LeuAspGlu-----LeuTyArgValThrArgGluTyThrGlnAsnLys 55
 DB 1289 CTGTCAAGCCTTCGTGATTTCTTCCCTACGAGTATTAAACACTACGTCGCGAAAGAG 1348
 QY 56 LysGluAlaGluLysIleLysAsnLeuIleLysThrValLleLysLeuAlaIleLeu 75
 DB 1349 GAGATGCCCTCATTAATTGTGATGATGCCGTGCTGATGATGATACATTGCT 1396
 QY 76 TyrArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLys 95
 DB 1397 -----TCTCTTAACAGATTTCAGATTGACAGTA-----CAGCATTCAC 1432
 QY 96 ValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyThrPheAspArg 115
 DB 1433 CTACATGAACCTGTTATCAAAACCTTGTCCAGACCAACCTTTTAT 1480
 QY 116 AsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArg 135
 DB 1481 -----ATGCTGCATCACTTCCTCGCAGTAC 1504
 QY 136 HisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCys 155
 DB 1505 CACGCTC-----CTCACGACCTCC 1522
 QY 156 GluPheLeuAlaAlaLeuTyArgAsnProPheGlyAsnPheLysProHisLeuGlnLysLeu 175
 DB 1523 AAACCTTGCTGCTGTGCTGTTATCCTAGAGAGTTTCTATCTCTGCTCATCAGCTA 1582
 QY 176 CysAspGlyIleAsnLysMetLeuAspGluGluAsn 187
 DB 1583 TCTCTGACATGCTGAAGCAGCTTTCAACAGCAAT 1618

RESULT 13
 US-09-016-434-1081
 ; Sequence 1081, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HERewith
 ; CLASSIFICATION:
 ; PRIORITY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1081:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1700 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: G1174071

US-09-016-434-1081

Alignment Scores:
 Pred. No.: 0.171 Length: 1700
 Score: 88.50 Matches: 41
 Percent Similarity: 42.27% Conservative: 41
 Best Local Similarity: 21.13% Mismatches: 56
 Query Match: 9.20% Indels: 56
 DB: 4 Gaps: 9

US-10-627-571-2 (1-188) x US-09-016-434-1081 (1-1700)

QY 16 GlnLysIleLeuGlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAsp 35
 DB 200 CAGCGAGGCACTTGTGAAGATGACTGTGAGTCCATCATGCGCTGCTGAGCGAG 259
 QY 36 ThrSerSerGlu-----ValLeuAspGluLeuTyArgValThrArgGluTyThr 52
 DB 260 GAGGCCAAGAGAGCCCGGCGGATCAACAGCAG-----ATCAGCGGAGCTCCGC 310
 QY 53 GlnAsnLysLysGluAlaGluLysIleLysAsnLeuIle----- 66
 DB 311 AGGACAAGCGGAGCGCGCGGAGCTCAAGCTGCTGCTCGCGGACAGAGAGAGT 370
 QY 67 -----LysThrValIleLys-----LeuAlaIleLeuTyArgAsnAsnGlnPheAsnGln 83
 DB 371 GGCAGAGTACCTTTATCAAGCATGAGATCATCTAGGCTCAGATCTCTATGATA 430
 QY 84 AspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrVal 103
 DB 431 GATTA-----AGGCGCTTCAACAGCTGCTGTATCAAGATCTTCAAGCGC 478
 QY 104 ValSerPheHis-----GlnValAspTyThrPheAspArgAsn 116


```

Db      479 ATGCAGCCATGATCAGACCATGACACTCAAGATCCATACAGTATGACACAAAT 538
Qy      117 ValLeuSerArgLeuLeuAenGluCyArgGluMetLeuIleGlnIleGlnArgHis 136
Db      538 ----- 538
Qy      137 LeuThrAlaLysSerHisGlyArg---ValAsnAsnValPheAspHisPheSerAspCys 155
Db      539 -----AAGGCTCATGCACAAATTAGTTGAGAAATT-----GATGTG 574
Qy      156 GluPheLeuAlaAlaLeuTyraAsnProPheGlyAsnPhelysProHisLeuGlnLys-Le 175
Db      575 GAGAAGAGTGTCTGTTTGAATCCATATGATGATGATGATGATGATGATGATGATGATGATGAT 634
Qy      175 uCyAspGlyIleAsnLysMetLeuAspGluGlnIle 188
Db      635 CTGGAATCCAGGAATCTATGATAGACGACGAGAAATATC 674

RESULT 14
US-09-023-655-898
; Sequence 898, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 898:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1174071
; US-09-023-655-898

Alignment Scores:
Pred. No.: 0.171 Length: 1700
Score: 88.50 Matches: 41
Percent Similarity: 42.27% Conservative: 41
Best Local Similarity: 21.13% Mismatches: 56
Query Match: 9.20% Indels: 56

```

```

DB:      4 Gaps:      9
US-10-627-571-2 (1-188) x US-09-023-655-898 (1-1700)
Qy      16 GluLysLysIleLeuGlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAsp 35
Db      200 CACGGAGGACGACCTTGGAAATAATGACTGCTGAGATCATCATGCGCTGCTGAGGAG 259
Qy      36 ThrSerSerGlu-----ValLeuAspGluLeuTyraArgValThrArgGluTyThr 52
Db      260 GAGGACCAAGGAAACCCGCGCGATCAACGAGAG-----ATCGACGCGAGCTCCG 310
Qy      53 GluAsnLysGluAlaGluLysLysIleLysAsnLeuLe----- 66
Db      311 AGGACACAGCGGAGCGCCGCCGAGCTCAAGCTGCTGCTGCTCGGACAGAGAGT 370
Qy      67 -----LysThrValIleLys---LeuAlaIleLeuTyraArgAsnGlnPheAsnGln 83
Db      371 GGCAAGATGACGCTTATCAAGCAGATGAGATCATCATGCGTCAGAGTACTGTGATGAA 430
Qy      84 AspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrVal 103
Db      431 GATATA-----AAGGCTTCAACCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 478
Qy      104 ValSerPheHis-----GluValAspTyThrPheAspArgAsn 116
Db      479 ATGCAGCCATGATCAGACCATGACACACTCAAGATCCATGATGATGATGATGATGATGAT 538
Qy      117 ValLeuSerArgLeuLeuAenGluCyArgGluMetLeuIleGlnIleGlnArgHis 136
Db      538 ----- 538
Qy      137 LeuThrAlaLysSerHisGlyArg---ValAsnAsnValPheAspHisPheSerAspCys 155
Db      539 -----AAGGCTCATGCACAAATTAGTTGAGAAATT-----GATGTG 574
Qy      156 GluPheLeuAlaAlaLeuTyraAsnProPheGlyAsnPhelysProHisLeuGlnLys-Le 175
Db      575 GAGAAGAGTGTCTGTTTGAATCCATATGATGATGATGATGATGATGATGATGATGATGATGAT 634
Qy      175 uCyAspGlyIleAsnLysMetLeuAspGluGlnIle 188
Db      635 CTGGAATCCAGGAATCTATGATAGACGACGAGAAATATC 674

RESULT 15
US-09-270-767-14175
; Sequence 14175, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14175
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
; US-09-270-767-14175

Alignment Scores:
Pred. No.: 0.196 Length: 1865
Score: 88.50 Matches: 41
Percent Similarity: 38.55% Conservative: 28
Best Local Similarity: 22.91% Mismatches: 67
Query Match: 9.20% Indels: 43
DB:      4 Gaps:      7
US-10-627-571-2 (1-188) x US-09-270-767-14175 (1-1865)

```


Oy		23	Mctvalseryserrisleaathrrhleuleapaspethrseerghlvalleusp	42
Db		652	TTGATGGATTAAAGACGATCATGAATATCTTCAAGAATGAGTAGA-----	607
Oy		43	GluDeuTykrvgtValThrArgslurTythrGlnasnLysLysgluaIagluLyslysile	62
Db		601	GAATATTTTCACGTACGACAC-----	589
Oy		63	LysasnLeuIllelysthrValIlleLysleuAlalle-----Leu	75
Db		580	-----ATAGAGCGACGATCATCGGTTTTAGGTTTGCTTACTTAGACAATAAGAAAATTGG	530
Oy		76	TyrArgasnAsnGlnPheasnGlnAspGluLeuAlaleMetGlnLysPheLysLysLys	95
Db		529	GAATAAGAAAAATGACAACTTAATAAAGAAATATCTCATTGAGAGAAAGAAATAAGAG	470
Oy		96	ValHisglnLeuAlamethThrValalaserPheHisglnValaspyrThrPheAsparg	115
Db		469	GTTCATTAATAGTTATTAACTTTAGACTTCCATTCATTAATAGAACACACATTTTAGAAC	411
Oy		116	AsnValleuserArg-----LeuleuasnGluCySarGluMetLeuHisglnlle	133
Db		409	AAGATGTATAATATTTTAGAACAAATATTAATTATATATGAATTTGATCATCAG---	355
Oy		133	IlglnIargHisleuthralalysSerHisglY-----	143
Db		352	CATGCCGCACACACAGAACTTAAGCATCATGCATCAAGCCATCAGCGTTGATTATAGA	299
Oy		144	-----ArgValasnAsnValPheasPhispheresasp-----	154
Db		292	AATACAAACGTTCTCGATTTTAGATTATCTTAACGGTATATTAATGCGCTGTTTTA	233
Oy		155	-----CysGluPheLeuAlalaLeuTYasnProPheGlyAsn	167
Db		232	GAAGCAAATGAGACGGTATAGCTGAACAAACAAACGACGATTCATATGATGTACC	173
Oy		168	PheLysProHisleuglnLys-----LeuCySaap-----GlyIlaAsnLysMetLeu	189
Db		172	GTATCATCCCTTGTTGCAAGAAAGCGCTTCATGACTTTTAGGAATTAACGAAATTA	113
Oy		184	AspGluGlu	186
Db		112	GATTAAGAA	104
<hr/>				
RESULT 18				
US-09-710-279-4029/c				
; Sequence 4029, Application US/09710279				
; Patent No. 6703492				
; GENERAL INFORMATION:				
; APPLICANT: KIMMERLY, WILLIAM JOHN				
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS				
; FILE REFERENCE: P03480US				
; CURRENT APPLICATION NUMBER: US/09/710,279				
; CURRENT FILING DATE: 2000-11-09				
; PRIOR APPLICATION NUMBER: 60/164,258				
; PRIOR FILING DATE: 1999-11-09				
; NUMBER OF SEQ ID NOS: 4472				
; SOFTWARE: Patentin Ver. 2.1				
; SEQ ID NO 4029				
; LENGTH: 3760				
; TYPE: DNA				
; ORGANISM: Artificial Sequence				
FEATURE:				
; OTHER INFORMATION: Description of Artificial Sequence: synthetic				
; OTHER INFORMATION: nucleic acid sequence				
US-09-710-279-4029				
<hr/>				
Alignment Scores:				
Pred. No.:	0.563	Length:	3760	
Score:	88.50	Matches:	46	
Percent Similarity:	39.41%	Conservative:	34	
Best Local Similarity:	22.66%	Mismatches:	64	

```

Query Match: 9.20% Indels: 59
DB: 4 Gaps: 9
US-10-627-571-2 (1-188) x US-09-710-279-4029 (1-3760)

QY 23 MetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAsp 42
   :::::||||| ::::: ::::: |||
DB 2847 TTGATGACTAAAGAAGTATGCATTAAGATATCTTACAAGTATAGATGGAGAA----- 2797
QY 43 GluLeuTyrArgValIThrArgGluTyrThrGlnAsnLysGluAlaGluLysIle 62
   ||| ::::: |||
DB 2796 GAATATTTTCCAGTCACGCCAC----- 2776
QY 63 LysAsnLeuIleLysThrValIleLysLeuAlaIle-----Leu 75
   ::::: ||| ||| :::::
DB 2775 -----AATAGAGCGAGTCATCGTTTATAGTGTCTTAACTTAGACAAATATGAAATTTG 2725
QY 76 TyrArgAsnAsnGlnPheAsnGlnAspGluLeuMetGluLysPheLysLysLys 95
   ::::: ||| :::::
DB 2724 GAAAGAAGAAATGAGAACTTAAAAAGAAATATCTCTATTAGAGAAAGAAATAAAGGAG 2665
QY 96 ValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyrThrPheAspArg 115
   ::::: ||| :::::
DB 2664 GTTCAATAAATGTTATTTAACTTTAGACTTTCCTATTCAAATAGACACACACTTTAGAAC 2605
QY 116 AsnValLeuSerArg-----LeuLeuAsnGluCysArgGluMetLeuHisGlnIle 132
   ::::: ||| :::::
DB 2604 AAGATGATAAATATATTTTATTTAGAACATATCTTATATTTATTAATGATTAAGATCATCAG-- 2548
QY 133 IleGlnArgHisLeuThrAlaLysSerHisGly----- 143
   ::::: |||
DB 2547 CATGCGCGACACAGAAACTTAAGCATCATGCACATCAAGCATCAGCTTGATTATAGAGA 2488
QY 144 -----ArgValAsnAsnValPheAspHisPheSerAsp----- 154
   ::::: ||| :::::
DB 2487 AATACAAACGTTTCTGCATTTTATGATTTATTTAAACGGTAAATATTAATGGCTTGTTTA 2428
QY 155 -----CysGluPheLeuAlaIleLeuTyrAsnProPheGlyAsn 167
   ||| |||
DB 2427 GGAGCAATGAGAGACGGTATAGCTGAACAAACAAACAGACAGATCATCAATGATGATAC 2368
QY 168 PheLysProHisLeuGlnLys-----LeuCysAsp-----GlyIleAsnLysMetLeu 183
   ||| ||| |||
DB 2367 GTACATCCCTTGTTGCAAGAAAGCGCTTCATGACTTTTATGGAATTAACAGAAATTA 2308
QY 184 AspGluGlu 186
   ||| |||
DB 2307 GATTAAGAA 2299

RESULT 19
US-09-616-289-46
; Sequence 46, Application US/09616289
; Patent No. 6632823
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Actona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,920
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1638)
US-09-616-289-46

Alignment Scores:
Pred. No.: 0.188 Length: 1638
Score: 88.00 Matches: 51
Percent Similarity: 36.28% Conservative: 31
Best Local Similarity: 22.57% Mismatches: 86
Query Match: 9.15% Indels: 58
DB: 4 Gaps: 10

US-10-627-571-2 (1-188) x US-09-616-289-46 (1-1638)

QY 14 GlnAlaGlnLysLysIleuGlyLysMetValSerLysSerIleAlaThr----- 31
Db 619 CAGAAAGCAATGAAGTCTCTACAGAAAAGCAGAGCCAGCTGTGCAAGAGAACACAC 678
QY 32 LeuIleAspAspThrSerSerGluValLeu-----Asp 42
Db 679 CTGCGCGGTGAGACAGCAAGCGCTCTGCGCCGACAGAACTTGAGAGCCATGCCGT 738
QY 43 GluLeuTyrrArgValThrArg-----GluTyrrGlnAsnLysGluVal 58
Db 739 GAGCTGCACCGGACCAACCGCTCTCCAAAGAAAGTGTCGACGCGCGCGGAGAGAG 798
QY 59 GluLysLysIleLysAsnLysLys-----ThrValIleLysLeu 72
Db 799 GAGGAAAGCGGACAGAGAGGTGACCTGCACCTTCAGGTGACCTGACATTCAGCTG 858
QY 73 AlaIleLeuTyrr-----ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMet 89
Db 859 CAGATGGAACAGACCAATGAGCCCACTCCAGCTGCCGCAAGAAACATGAGCTGGCT 918
QY 90 GluLysPheLysLysValHisGlnLeuAlaMetThrValAlaSerPheHisGlnVal 109
Db 919 GAGAGGCTCAAGAAAGCTGATTGAGCATGTGAGCTCGCGAG-----GAGCATATTC 969
QY 110 AspTyrrThrPhe-----AspArgAsnValLeuSerArgLeuLeuAsnGln 124
Db 970 GACAAAGTCTTCAAAACAAAGACCTTACAAACAGCTGTGGATGCCAAGCTCCAGCAG 1029
QY 125 CysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArg 144
Db 1030 GCCCAGAGATGCTTAAGGACGCAAGAACGGCGCACAGCGGAGAAAGATTTT---CTC 1086
QY 145 ValAsnAsnValPheAspHisPheSerAspCysGlnPheLeu----- 158
Db 1087 CTGAAAGAGCACTAAGTCCCGACAGAGATGTGTGAGTATGAAAGACAGAACCCAC 1146
QY 159 -----AlaAlaLeuTyrrAsnPro----- 164
Db 1147 CTGAAGCAACAGCTTCCCTATACACAGAGAACTTGAGAGAGTTCCAGAAACACATTTCC 1206
QY 165 -----PheGlnAsnPheLysProHisLeuGlnLysLeuCysAspGlyIle 179
Db 1207 AAAGCAGGAGGTATTTCACACATTCACAGCAGAGATGGAAGAAAGTACTAAGAAAGATC 1266
QY 180 AsnLysMetLeuAspGln 185
Db 1267 AAGAACTGAGAAAGAA 1284

RESULT 20
; Sequence 3, Application US/09750590A
; Patent No. 6780987
; GENERAL INFORMATION:

```

```

; APPLICANT: Herman, Ira
; APPLICANT: Welch, Alice
; TITLE OF INVENTION: BETA-CAP73 CONTROL OF NORMAL AND
; TITLE OF INVENTION: ABNORMAL CELL MIGRATION
; FILE REFERENCE: TUI-001CP
; CURRENT APPLICATION NUMBER: US/09/750,590A
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/170,182
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: 09/733,818
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4206
; TYPE: DNA
; ORGANISM: Bos taurus
US-09-750-590A-3

Alignment Scores:
Pred. No.: 0.775 Length: 4206
Score: 88.00 Matches: 36
Percent Similarity: 43.46% Conservative: 47
Best Local Similarity: 18.85% Mismatches: 70
Query Match: 9.15% Indels: 38
DB: 4 Gaps: 7

US-10-627-571-2 (1-188) x US-09-750-590A-3 (1-4206)

QY 9 LysAsnLeuAlaValGlnAlaGlnLysLysIleuGlyLysMetValSerLysSerIle 28
Db 2176 AAAAATGCCCTTAAAGTAAGTGAAGAAATG-----AAAAAGTCA 2217
QY 29 AlaThrThrLeuIleAspAspThrSerSerGluValIleAspGlnLeuTyrrArgValThr 48
Db 2218 CATGATGTAAATGTTGATGATGATTGAATAAAAGCTTCAGAT-----GTGACA 2265
QY 49 ArgGluTyrrThrGlnAsnLysLysGluAlaGlnLysLysIle-----LysAsnLeu 65
Db 2266 CACAAATATACAGAAAGAAAGTTGAATGAGAAAGTTGCTTATGAAAAATGCCAGTTTA 2325
QY 66 IleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnAsnGlnPheAsnGlnAspGln 85
Db 2326 AGTAAATAATGTCAGCGCGCTGAAAACGTGTTCAATACCTCCGAGAGACAGAAAGAA 2385
QY 86 LeuAlaLeuMetGln-----LysPheLysLysValHisGlnLeuAlaMet 101
Db 2386 ATGATGGCTCTGAAATCCAAATATCACTGAACCTTAAAGACAGCTGTCTGAACCTTAATAA 2445
QY 102 ThrValAlaSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArgLeu 121
Db 2446 AAATGTGTGAA-----GACCAAGAGAAATATATTCACATC 2481
QY 122 LeuAsnGlnCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
Db 2482 ATGCTGAAACAAATGATTTGAAGAAAGACCAAGATGATCACTAGTATGTCGCCGTAAGAAC 2541
QY 142 HisGlnArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAlaLeu 161
Db 2542 CATGAAGATTAATAAATGCC-----TTGAGTGCACA 2574
QY 162 TyrrAsnProPheGlnAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
Db 2575 TTGATTAATAAACAATATGAGAAATTAAGTATGTAAGAAAGATGTAATATTAATCA 2634
QY 181 -----LysMetLeuAspGlnLysnle 188
Db 2635 GAATTTGTGAATAATAAAGATGAGAACGAAATA 2667

RESULT 21
; Sequence 17, Application US/08979608A
; Patent No. 6355451

```



```

; LENGTH: 4697 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 3...1592
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-517-849-17

Alignment Scores:
Pred. No.: 0.915 Length: 4697
Score: 88.00 Matches: 51
Percent Similarity: 36.28% Conservative: 31
Best Local Similarity: 22.57% Mismatches: 86
Query Match: 9.15% Indels: 58
DB: 4 Gaps: 10

US-10-627-571-2 (1-188) x US-09-517-849-17 (1-4697)
QY 14 GlnAlaGlnLysLysIleuGlnLysMetValSerLysSerIleAlaThr----- 31
Db 573 CAGAACGATGATGATGCTCTACAGAAAAGCAGAGCCAGCTGTGCAGAGAAAGACAC 632
QY 32 LeuIleAspAspThrSerSerGluValLeu-----Asp 42
Db 633 CTGCGGGGTGAGCAGCAGCAAGCGCGTCTGCGCCGACGACCTTGAGAGCTTATGCCGT 692
QY 43 GluLeuTyArgValThrArg-----GluTyThrGlnAsnLysGluAla 58
Db 693 GAGCTCAGCGGACCAACCGCTCCCTCAAGAGAAGGTGTGACGGCGCGGAGAGAG 752
QY 59 GluLysLysIleLysAsnLeuIleLys-----ThrValIleLysLeu 72
Db 753 GAGGAGAGAGGCAAGAGAGGTGACCTCGCAGCTTCCAGGTGACCTGAATGACATTGAGCTG 812
QY 73 AlaIleLeuTyR-----ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMet 89
Db 813 CAGATGGAACAGCACAATGAGCGCAACTCCAACTGCGCCAAAGAAATGAGAGCTGGCT 872
QY 90 GluLysPheLysLysValHisGlnLeuAlaMetThrValValSerPheHisGlnVal 109
Db 873 GAGAGGCTCAAGAGAGCTGATTGAGCAGTATGAGACTGCGCGAG-----GAGCATATC 923
QY 110 AspTyThrPhe-----AspArgAsnValLeuSerArgLeuLeuAsnGln 124
Db 924 GACAAAGTCTTCAAAACAAGACCTTACACACAGCAGCTGGTGAATGCCAAAGCTCCAGCAG 983
QY 125 CysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlnArg 144
Db 984 GCCCAGAGATGCTTAAAGAGGAGGCAAGAGCGGACACAGCGGAGAGAAAGATTTT---CTC 1040
QY 145 ValAsnAsnValPheAspHisPheSerAspCysGluPheLeu----- 158
Db 1041 CTGAAGAGGAGCAGTAAGTCCAGAGAGATGTGTGAGCTGTGAAGCAGCAAGACCCAC 1100
QY 159 -----AlaIleLeuTyArgAsnPro----- 164
Db 1101 CTGAAGCAACAGCTTCTATACACAGAGAGAGTTTGAGAGTTCCAGAAACACACTTTC 1160
QY 165 -----PheGlnAsnPheLysProHisLeuGlnLysLeuCysAspGlnIle 179
Db 1161 AAAAGCAGGAGATATTTCACCAATTCACAGCAGAGATGTGAAAAGATGACTAAGAAAGATC 1220
QY 180 AsnLysMetLeuAspGln 185
Db 1221 AAGAAAGCTGAGAAAGAA 1238

RESULT 23
US-09-616-289-17
; Sequence 17, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:

```

```

; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIORITY FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-09-616-289-17

Alignment Scores:
Pred. No.: 0.915 Length: 4697
Score: 88.00 Matches: 51
Percent Similarity: 36.28% Conservative: 31
Best Local Similarity: 22.57% Mismatches: 86
Query Match: 9.15% Indels: 58
DB: 4 Gaps: 10

US-10-627-571-2 (1-188) x US-09-616-289-17 (1-4697)
QY 14 GlnAlaGlnLysLysIleuGlnLysMetValSerLysSerIleAlaThr----- 31
Db 573 CAGAACGATGATGATGCTCTACAGAAAAGCAGAGCCAGCTGTGCAGAGAAAGACAC 632
QY 32 LeuIleAspAspThrSerSerGluValLeu-----Asp 42
Db 633 CTGCGGGGTGAGCAGCAGCAAGCGCGTCTGCGCCGACGACCTTGAGAGCTTATGCCGT 692
QY 43 GluLeuTyArgValThrArg-----GluTyThrGlnAsnLysGluAla 58
Db 693 GAGCTCAGCGGACCAACCGCTCCCTCAAGAGAAGGTGTGACGGCGCGGAGAGAG 752
QY 73 AlaIleLeuTyR-----ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMet 89
Db 753 GAGGAGAAACGCAAGAGAGGTGACCTCGCAGCTTCCAGGTGACACTGAATGACATTGAGCTG 812
QY 73 AlaIleLeuTyR-----ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMet 89
Db 813 CAGATGGAACAGCACAATGAGCGCAACTCCAACTGCGCCAAAGAAATGAGAGCTGGCT 872
QY 90 GluLysPheLysLysValHisGlnLeuAlaMetThrValValSerPheHisGlnVal 109
Db 873 GAGAGGCTCAAGAGAGCTGATTGAGCAGTATGAGCTGCGGAG-----GAGCATATC 923
QY 110 AspTyThrPhe-----AspArgAsnValLeuSerArgLeuLeuAsnGln 124
Db 924 GACAAAGTCTTCAAAACAAGACCTTACACACAGCAGCTGGTGAATGCCAAAGCTCCAGCAG 983
QY 125 CysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlnArg 144
Db 984 GCCCAGAGATGCTTAAAGAGGAGGCAAGAGCGGACACAGCGGAGAGAAAGATTTT---CTC 1040
QY 145 ValAsnAsnValPheAspHisPheSerAspCysGluPheLeu----- 158
Db 1221 AAGAAAGCTGAGAAAGAA 1238

```

Db 1041 CTGAAGAGGACGTAGAGTCCACAGAGATGTGAGCTGATGAAGCAGCAAGACCAC 1100
 Qy 159 -----AlaIaLeuTYrAsnPro----- 164
 Db 1101 CTGAAGCAGACGTTGCTTATACACAGAGAAGTTTGAGAGTTCCAGAAACACTTTC 1160
 Qy 165 -----PheGlyAsnPhelyProHisLeuGlnIleuGlySerIleuGlyVal 179
 Db 1161 AAAAGCAGCAGATATTCACCACTTCAGCAGAGAGATGAGAAAGATGACTTAAGAGATC 1220
 Qy 180 AsnIleuMetLeuAspGlu 185
 Db 1221 AAGAGCTGAGAAAGAA 1238

RESULT 24
 US-09-750-590A-1
 ; Sequence 1, Application US/09750590A
 ; Patent No. 6780987
 ; GENERAL INFORMATION:
 ; APPLICANT: Herman, Ira
 ; APPLICANT: Welch, Alice
 ; TITLE OF INVENTION: BETA-CAP73 CONTROL OF NORMAL AND
 ; TITLE OF INVENTION: ABNORMAL CELL MIGRATION
 ; FILE REFERENCE: TUI-001CP
 ; CURRENT APPLICATION NUMBER: US/09/750,590A
 ; CURRENT FILING DATE: 2000-12-28
 ; PRIOR APPLICATION NUMBER: 60/170,182
 ; PRIOR FILING DATE: 1999-12-10
 ; PRIOR APPLICATION NUMBER: 09/733,818
 ; PRIOR FILING DATE: 2000-12-08
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 4730
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (392)...(4597)
 ; US-09-750-590A-1

Alignment Scores:
 Pred. No.: 0.925 Length: 4730
 Score: 88.00 Matches: 36
 Percent Similarity: 43.46% Conservative: 47
 Best Local Similarity: 18.85% Mismatches: 70
 Query Match: 9.15% Indels: 38
 DB: 4 Gaps: 7

US-10-627-571-2 (1-188) x US-09-750-590A-1 (1-4730)
 Qy 9 LysAsnLeuAlaValGlnIleuGlnIleuGlyLeuGlyMetValSerIle 28
 Db 2567 AAAAATGTCCTTAAAGTAAGTAAGAAATG-----AAAAGTCA 2608
 Qy 29 AlaThrThrLeuIleAspThrSerSerGluValLeuAspGluLeuTYrArgValThr 48
 Db 2609 CATGATGATATGTTGATGATTTGAATTAAGACTTCACAT-----GTGACA 2656
 Qy 49 ArgGluTYrThrGlnAsnIleuGlyValGlnIleuGlyLeu-----LysAsnLeu 65
 Db 2657 CACAAATATACAAAAGAAAGTTGAAATGGAAGAGTGTCTTATGAAAATGCCAGTTTA 2716
 Qy 66 IleLeuThrValIleLysLeuAlaIleLeuTYrArgAsnAsnGlnPheAsnGlnAspGlu 85
 Db 2717 AGTAATAATGTCACGCCCTGAAACGTGTTCAATCCCGAGAGACAGCAAAAAGAA 2776
 Qy 86 LeuAlaLeuMetGlu-----LysPheIleLysIleValHisGlnLeuAlaMet 101
 Db 2777 ATGATGCTCTGAAATCCAAATATACACTGAACTTAAGAGACGCTGCTGAACTTAATAA 2836
 Qy 102 ThrValValSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArgLeu 121

Db 2837 AATATGTGTGA-----GACCAAGAGAAAATATATATCACTC 2872
 Qy 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
 Db 2873 ATGTCTGAAAACAAATGATTTGAAAAAGACCATGATGATCAGATATGCCCCGTGAAAACC 2932
 Qy 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleu 161
 Db 2933 CATGAAGAGATTAAAACTGCC-----TTGAGTACGACA 2965
 Qy 162 TYrAsnProPheGlyAsnPhelyProHisLeuGlnIleuGlyCysAspGlyIleAsn--- 180
 Db 2966 TTGATTAATAACCAATAGAGATTAAGATTAAGATTAAGAGAGAGTGAATTAATCA 3025
 Qy 181 -----LysMetLeuAspGluIleuAsnIle 188
 Db 3026 GAATTTGTAATAATTAAGATGAGAACGAAATTA 3058

RESULT 25
 US-09-976-594-940
 ; Sequence 940, Application US/09976594
 ; Patent No. 6673549
 ; GENERAL INFORMATION:
 ; APPLICANT: Furness, Michael
 ; APPLICANT: Buchinder, Jenny
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 ; FILE REFERENCE: PA-0041 US
 ; CURRENT APPLICATION NUMBER: US/09/976,594
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,409
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 1143
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 940
 ; LENGTH: 3172
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. 6673549 481154.3
 ; US-09-976-594-940

Alignment Scores:
 Pred. No.: 0.59 Length: 3172
 Score: 87.50 Matches: 44
 Percent Similarity: 42.18% Conservative: 45
 Best Local Similarity: 20.85% Mismatches: 87
 Query Match: 9.10% Indels: 35
 DB: 4 Gaps: 8

US-10-627-571-2 (1-188) x US-09-976-594-940 (1-3172)
 Qy 9 LysAsnLeuAlaValGlnIleuGlnIleuGlyLeuGlyMetValSerIle 28
 Db 1140 AGGAATTTG-----GCCCAACGAAATTAATTAAGAAATGAGTCTTAAGTATTA 1190
 Qy 29 AlaThrThrLeuIleAspThr-----SerSerGluValLeuAspGluLeu 44
 Db 1191 GAACAACAATTCGACAGAAAACAAGCTTTAAAGACACAGAAAACATGAAGAGCTTA 1250
 Qy 45 TYrArgValThrArgGluTYrThrGlnAsnLys--LysGluAlaGluLysIleLys 63
 Db 1251 GTATGCAACCTTCCTCCGATGACTCAAAATCAAAATTAAGAAAGAGAAAGATCCAG 1310
 Qy 64 AsnLeuIleLysThrValIleLysLeuAlaIleLeuTYrArgAsnAsnGlnPheAsnGln 83
 Db 1311 GATTTCTGAAGCTCAGCAAAAATATACCAACATTTGTTAAAGAAATGAAGCAAAAGAT 1370
 Qy 84 AspGluLeuAlaLeuMetGluLysPheLysIleValHisGlnLeuAlaMetThrVal 103
 Db 1371 CTTGAATACAGATACACAAGAAAGAAATGAAATTTATTCGAGACTGAGAGAGATT 1430
 Qy 104 ValSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArgLeuLeuAsn 123

```

Db      1431 GCTAACTGATGACAC---ATTGGAATGAAAGAAACAAATTGTTAACTACTCCAC 1487
      ::::|||||
      124 GluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeu----- 137
      ::::|||||
      1488 AAAAGCTCATGAAAGAAATGAAATGAAAGAAAGGATGATTAATTAATGAACCTT 1547
      ::::|||||
      138 -----ThraIalysSerHisGlyArg-----ValAsnAsnValPheAsp 150
      ::::|||||
      1548 GAAATTCTGAGAAATAGTGGCCGTTAGTCAAGAAAGAAAGCTACAAATTCATGCTGAAA 1607
      ::::|||||
      151 HisPheSerAspCysGluPheLeuAlaIleuTyraAsnProPheGlyAsnPhelysPro 170
      ::::|||||
      1608 CACGCCACACATGTTACATCAAGAGAGCATCAAAACGATGTGCGCAAAATTGTATCA 1667
      ::::|||||
      171 HisLeuGlnHisLeu-----CysAspGlyIleAsnLys----- 181
      ::::|||||
      1668 AAACCTTCAGGAATGAAAGAAAGAAAGAAAGCCAGTTAAATTAACATTGACAGACTTGCC 1727
      ::::|||||
      182 -----MetLeuAspGluGluAsnIle 188
      ::::|||||
      1728 AACACGATCAACATGATCGAAGAGAGATGCTG 1760
      ::::|||||

RESULT 26
US-09-336-447A-2
; Sequence 2, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP21 AND USP22 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3349
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-2

Alignment Scores:
Pred. No.: 0.64 Length: 3349
Score: 87.50 Matches: 47
Percent Similarity: 40.36% Conservative: 43
Best Local Similarity: 21.08% Mismatches: 66
Query Match: 9.10% Indels: 67
DB: Gaps: 10

US-10-627-571-2 (1-188) x US-09-336-447A-2 (1-3349)
QY      4 AspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMet 23
      :|||:|||||
      1296 GACATTCTTAACACCAAGATGATCAAGATCTTCAGAGGAGGTAAGGT----- 1349
      :|||:|||||
      24 ValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal----- 40
      :|||:|||||
      1350 -----CTTGATATAGAGGTGGGTGAATTAAACCGAGAC 1382
      :|||:|||||
      41 LeuAspGluLeuTyraArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGluLys 60
      :|||:|||||
      1383 ATTAATTCACCTTCATATGTTACTGCAACCAACCAAGATCAACAAAGGCTTAAGAG 1442
      :|||:|||||
      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAla-----IleLeuTy 76
      :|||:|||||
      1443 GGGGTAAAGAGCTGATTAATGAGGTGGGTGATTAAGCCGAGACATTAACTCACTTCAT 1502
      :|||:|||||
      77 ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeu----- 88
      :|||:|||||

```

```

Db      1503 GATGATGTTGCTGACACCAAGATGACATTGCTAAACCAAGCTGACATCAAGGCTTT 1562
      :|||:|||||
      89 -----MetGluLysPheLysLysValIleIleGlnLeuAlaMetThrValSer 105
      :|||:|||||
      1563 AATAAGAGAGTGAAGAGAGCTTGATAGAGAGGTGGGTGATTAAGCCGAGACATTGGTTCA 1622
      :|||:|||||
      106 PheHis-----GlnValAspTy----- 111
      :|||:|||||
      1623 CTTGATGATGATGTTGCCACCAACGATGACATTGCTTAAACCAAGCGGATATCAA 1682
      :|||:|||||
      112 ThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGln 131
      :|||:|||||
      1683 ACACCTGAACCAATGTCGAGAGAAATTAATAATCTAAAGCGGTGCGCTGCTGATCAG 1742
      :|||:|||||
      132 IleIleGlnArgHisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAsp--- 150
      :|||:|||||
      1743 -----AAAGCGGATATGTTGATTAATTAACATCAACAAATTAATGAGCTG 1784
      :|||:|||||
      151 -----HisPheSerAspCysGluPheLeuAlaIleuTyraAsnPro 164
      :|||:|||||
      1785 GCACAAACGACAGATCAGATAGCTCTGATACAAACACTT----- 1826
      :|||:|||||
      165 PheGlyAsnPhelysProHisLeuGlnLys---LeuCysAspGlyIleAsnLysMetLeu 183
      :|||:|||||
      1827 -----AAAAACAATGTCGAGAGAGGTTATTGATCTTAAGCGGTGCTCATTT 1874
      :|||:|||||
      184 AspGluGlu 186
      :|||:|||||
      1875 GATCAAAA 1883
      :|||:|||||

RESULT 27
US-09-952-267B-2
; Sequence 2, Application US/0995267B
; Patent No. 6753417
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP21 AND USP22 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/952,267B
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US/09/336,447
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3349
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-09-952-267B-2

Alignment Scores:
Pred. No.: 0.64 Length: 3349
Score: 87.50 Matches: 47
Percent Similarity: 40.36% Conservative: 43
Best Local Similarity: 21.08% Mismatches: 66
Query Match: 9.10% Indels: 67
DB: Gaps: 10

US-10-627-571-2 (1-188) x US-09-952-267B-2 (1-3349)
QY      4 AspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMet 23
      :|||:|||||
      1296 GACATTCTTAACACCAAGATGATCAAGATCTTCAGAGGAGGTAAGGT----- 1349
      :|||:|||||
      24 ValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal----- 40
      :|||:|||||
      1350 -----CTTGATATAGAGGTGGGTGAATTAAACCGAGAC 1382
      :|||:|||||

```



```

RESULT 29
US-09-517-849-14
: Sequence 14, Application US/09517849
: Patent No. 6605368
: GENERAL INFORMATION:
: APPLICANT: Lees, Ann M.
:               Lees, Robert S.
:               Law, Simon W.
:               Arjona, Anibal A.
: TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
:                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
:                   TREATING ATHEROSCLEROSIS
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/517,849
: FILING DATE: 02-Mar-2000
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/979,608
: FILING DATE: 26-NOV-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Myers, Louis
: REGISTRATION NUMBER: 35,965
: REFERENCE/DOCKET NUMBER: 10797-003001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4722 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: Coding Sequence
: LOCATION: 61..1731
: SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-517-849-14

Alignment Scores:
Pred. No.:          1.07          Length:      4722
Score:              87.50          Matches:     42
Percent Similarity: 36.79%         Cons:       36
Best Local Similarity: 19.81%      Mismatches:  91
Query Match:        9.10%         Indels:     43
DB:                 4             Gaps:        7

US-10-627-571-2 (1-188) x US-09-517-849-14 (1-4722)
Oy      13  ValGlnAlaGluLysLeuGlyLysMetValSerLysSerIleAlaThrLeu 32
Db      721 GtGcGAGGAGAAAGACCACTGCGTGGCGAGCAACAGAGCCATCTGCGCCGACAGAG 780
Oy      33  IleApAspThrSerSerGluValIleuAspGluLeuTyArgValThrArgGluTyThr 52
Db      781 CTCGAGAGCCTGTGCCGAGGCTGCAGCGCGCACCAACCGCTCGCTCAAGAGAAAGGTGTG 840
Oy      53  GlnAenLysLysGlnAlaGluLysLysIleLysAsnLeuLys----- 67
Db      841 CAGCGAGCCCGAGAGGAGAGAGAAAGCCGACAGAGGTGACGTCACTTCACAGATGACG 900

```

[illegible]

```

RESULT 30
US-09-616-289-14
Sequence 14, Application US/09616289
Patent No. 6632923
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
CURRENT FILING DATE: 2000-07-14
PRIORITY APPLICATION NUMBER: US 09/517,849
PRIORITY FILING DATE: 2000-03-02
PRIORITY APPLICATION NUMBER: US 08/979,608
PRIORITY FILING DATE: 1997-11-26
PRIORITY APPLICATION NUMBER: US 60/031,930
PRIORITY FILING DATE: 1996-11-27
PRIORITY APPLICATION NUMBER: US 60/048,547
PRIORITY FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 4722
TYPE: DNA
ORGANISM: Oryctolagus cuniculus
FEATURE:
NAME/KEY: CDS
LOCATION: (61)...(1731)
US-09-616-289-14

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

1.07
87.50
36.78%
19.81%
9.10%

Length: 4722
Matches: 42
Conservative: 36
Mismatches: 91
Indels: 43
Gaps: 7

```


TELEFAX: 201-753-6237
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2651 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-462-949-1

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.708	86.00	45	32	50	96
Percent Similarity:	34.53%					
Best Local Similarity:	20.18%					
Query Match:	8.94%					

US-10-627-571-2 (1-188) x US-08-462-949-1 (1-2651)

```

QY 31 ThrLeuIlAspAspThrSerSerGluValLeuAspGluLeuTyArgValThrArgGlu 50
    |||||
DB 1768 ACTCTCCGCCCGGATGGTTCCTTATGATCTGAATATCTTTTCATATAGCCAGTCC 1709
    |||||
QY 51 TyrThrGln-----AsnLysLysGluAlaGluLysLysIleLys 63
    |||||
DB 1708 CAGACCGGAACCTTTTACATCGTCAGATTACCTGTATGACACCGCATAGAACCTTATCAA 1649
    |||||
QY 64 AsnLeuIleLysThrValIle-----LysLeuAlaIleLeuTyArgAsnAsnGln 80
    |||||
DB 1648 CACCTTCTTCAGGTTTCTTTTAAATACCAATACGGTGTCTGTATCAATATCATCATG 1589
    |||||
QY 81 Phe-----AsnGlnAspGluLeuAlaLeuMetGlu 90
    |||||
DB 1588 TTTACCGGGTTCGTTCCAGGTTCTCAGTCGACGACACAGGTCCTTCGGGCTTCTT--- 1532
    |||||
QY 91 LysPheLysLysValHisGlnLeuAla----- 100
    |||||
DB 1531 GATATTCAACCGGTTATCCATCAGTTGAATATCACTTGTCATATGCAATTCGTCGCC 1472
    |||||
QY 101 -----MetThrValIleSerPheHis 107
    |||||
DB 1471 AATATATACAGAGATTGGTAGCCTCAAGTTCGATGTATTTACCATATAGGCTTCAC 1412
    |||||
QY 107 ----- 107
    |||||
DB 1411 CTTGAGAGCGAAGTATGTATTAACCCCGAATCAACTCATTTGGCATTCGATCTTTCAT 1352
    |||||
QY 108 -----GlnValAspTyThrPheAspArgAsnValLeuSerArg--- 120
    |||||
DB 1351 CTTCACTGTGCTTCTGTCTGTGCAACATGACATCTGAAGAATCTTCTCCACCTGGAA 1292
    |||||
QY 121 -----LeuLeuAsnGluCysArgGluMetLeuHisGlnIle----- 132
    |||||
DB 1291 TCTCAACCGCTGCTTTCATCATTAATGACTCCCGATATATGCTGGAACATCTCGCGTGA 1232
    |||||
QY 133 -----LysGlnArgHisLeuThrAlaLysSerHisGlyArg----- 144
    |||||
DB 1231 TTTCCGGAATGCGGATCTTTTCAGCATCTGCGCGGACGAGATGGTATATCTCGGTG 1172
    |||||
QY 145 -----ValAsnAsnValPheAspHisPheSerArgPheCysGluPhe 157
    |||||
DB 1171 CTTGAACCTGCTGTGATATGTGTGAATAAGTCAATCTTCATTTTC-----TGTGAGCTT 1118
    |||||
QY 158 LeuAlaAla 160
    |||||
DB 1117 CTTCTGGCG 1109
    |||||

```

RESULT 37
 US-08-023-764B-1/c
 Sequence 1, Application US/08023764B

Patent No. 5679540
 GENERAL INFORMATION:
 APPLICANT: Rasmussen, Beth Ann
 TITLE OF INVENTION: Cloning and identification of a two
 TITLE OF INVENTION: Component Signal Transducing Regulatory System from
 TITLE OF INVENTION: Bacteroides Fragilis
 NUMBER OF SEQUENCES: 39
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: American Cyanamid Company
 STREET: One Cyanamid Plaza
 CITY: Wayne
 STATE: New Jersey
 COUNTRY: United States
 ZIP: 07470-8426
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/023,764B
 FILING DATE: 26-FEB-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Barnhard, Elizabeth M.
 REGISTRATION NUMBER: 31,088
 REFERENCE/DOCKET NUMBER: 31,658-00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (201)831-3246
 TELEFAX: (201)831-3305
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2651 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-023-764B-1

Alignment Scores:
 Pred. No.: 0.708 Length: 2651
 Score: 86.00 Matches: 45
 Percent Similarity: 34.53% Conservative: 32
 Best Local Similarity: 20.18% Mismatches: 50
 Query Match: 8.94% Indels: 96
 DB: 1 Gaps: 10

US-10-627-571-2 (1-188) x US-08-023-764B-1 (1-2651)

```

QY 31 ThrLeuIlAspAspThrSerSerGluValLeuAspGluLeuTyArgValThrArgGlu 50
    |||||
DB 1768 ACTCTCCGCCCGGATGGTTCCTTATGATCTGAATATCTTTTCATATAGCCAGTCC 1709
    |||||
QY 51 TyrThrGln-----AsnLysLysGluAlaGluLysLysIleLys 63
    |||||
DB 1708 CAGACCGGAACCTTTTACATCGTCAGATTACCTGTATGACACCGCATAGAACCTTATCAA 1649
    |||||
QY 64 AsnLeuIleLysThrValIle-----LysLeuAlaIleLeuTyArgAsnAsnGln 80
    |||||
DB 1648 CACCTTCTTCAGGTTTCTTTTAAATACCAATACCGTGTCTGTATGGAATCATCATG 1589
    |||||
QY 81 Phe-----AsnGlnAspGluLeuAlaLeuMetGlu 90
    |||||
DB 1588 TTTACCGGGTTCGTTCCAGGTTCTCAGTCGACGACACAGGTCCTTCTTCGGGCTTCTT--- 1532
    |||||
QY 91 LysPheLysLysValHisGlnLeuAla----- 100
    |||||
DB 1531 GATATTCAACCGGTTATCCATCAGTTGAATATCACTTGTCATATGCAATTCGTCGCC 1472
    |||||
QY 101 -----MetThrValIleSerPheHis 107
    |||||

```

Db 1471 AATATTAACAGATGTGAGCTCAAGTTCGATGTATCTTACATTATAGGCTTCAC 1412
 Qy 107 ----- 107
 Db 1411 CTTAGAGCGAAGTATGTATTAACCCGAAATCACTGATGGATCGAGTCTTTTCAT 1352
 Qy 108 -----GlnValAspTyr-ThrPheAspArgAsnValLeuSerArg--- 120
 Db 1351 CTTACAGTGTCTTCTGTCTGTGACATAGACATGTGAAGAATTCTTCACCTGGAA 1292
 Qy 121 -----LeuLeuAsnGluCysArgGluMetLeuHisGlnIle----- 132
 Db 1291 TCTCAACCGCTTGTTTCATCATTAATGACTCCGATATATGCTCGAACAATCTGGGTGA 1232
 Qy 133 -----IleGlnArgHisLeuThrAlaLysSerHisGlyArg----- 144
 Db 1231 TTTCCGGAATCGGAGATCTTTCAGACATCTGCGCGCAAGAGATGGATATCGGTGT 1172
 Qy 145 -----ValAsnAsnValPheAspHisPheSerAspCysGluPhe 157
 Db 1171 CTTGAACCTGCTGTCTCATTTGTTGATTAAGTCACTTCTCATTTTC-----TGTCAAGCTT 1118
 Qy 158 LeuAlaAla 160
 Db 1117 CTTCTGCGCG 1109

RESULT 38
 US-09-134-000C-1489
 ; Sequence 1489, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1489
 ; LENGTH: 618
 ; TYPE: DNA
 ; ORGANISM: Enterococcus faecalis
 ; US-09-134-000C-1489

Alignment Scores:
 Pred. No.: 0.0922 Length: 618
 Score: 85.50 Matches: 39
 Percent Similarity: 43.21% Conservative: 31
 Best Local Similarity: 24.07% Mismatches: 52
 Query Match: 8.89% Indels: 41
 DB: 4 Gaps: 6

US-10-627-571-2 (1-188) x US-09-134-000C-1489 (1-618)
 Qy 5 ValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMetVal 24
 Db 25 ATTATCAACGACGACATTTTGTCTTCAAAATCAAAAACCAATGAAGG----- 75
 Qy 25 SerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeu 44
 Db 76 -----ATTACGGTTGACTTAGATACATTAACGTTGAATTACTAGAGAAAGA 123
 Qy 45 TyrArgValThrArgGlu-----TyrThrGlnAsn 54
 Db 124 TATGAATTTTCAGAGAGAAAAATCTGATATTCTCGAAAAAGTATGAGCAAGCG 183
 Qy 55 LysLysGluAlaGluLysLysIleLysAsnLeuIleLysThrValIleLysLeuAlaIle 74
 Db 184 GCACCTTCATCGAAGAAATATCAAAAGAAATTAAGAGCAATCAAAAGCAATT----- 237

Qy 75 LeuTyrArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLys 94
 Db 238 -----GGAAGAAATTTAACAAGATGATGTTGAATTAATGGCTTCTGTAAGAA 288
 Qy 95 LysValHisGlnLeuAlaMetThrValValIleSerPheHisGlnValAspTyrThrPheAsp 114
 Db 289 GCGGTG-----TATGATAGCATTTAGAGAAAGCG--GGA 319
 Qy 115 ArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIle----- 132
 Db 320 CGATTGACATGAGCAGATTGCTGACCAAGTGTTAAGAAATGTATCAGTCAATAATTAG 379
 Qy 133 -----IleGlnArgHisLeuThrAlaLysSerHisGlyArgValAsn 146
 Db 380 CTTACCAAGAGAGATGAATGAATGAAATGTTGTGCACAAAA--CACCGCGCATGACAG 436
 Qy 147 AsnVal 148
 Db 437 AATCTC 442

RESULT 39
 PCT-US94-01101-1/C
 ; Sequence 1, Application PC/TUS9401101
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; APPLICANT:
 ; APPLICANT:
 ; APPLICANT:
 ; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT B-CELL GROWTH
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: WOLF, GREENFIELD & SACKS, P.C.
 ; STREET: 600 ATLANTIC AVENUE
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/01101
 ; FILING DATE: FILED HEREWITH
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/005,156
 ; FILING DATE: 15-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: GATES, EDWARD R.
 ; REGISTRATION NUMBER: 31,616
 ; REFERENCE/DOCKET NUMBER: B0819/7000WO
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/720-3500
 ; TELEFAX: 617/720-2441
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1854 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: HOMO SAPIENS
 ; CELL TYPE: LYMPHOMA
 ; PCT-US94-01101-1

Alignment Scores:

[illegible]

```

; SEQ ID NO 1541
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1541

Alignment Scores:
Pred. No.: 0.951 Length: 2637
Score: 85.00 Matches: 37
Percent Similarity: 48.52% Conservative: 45
Best Local Similarity: 21.89% Identical: 37
Query Match: 8.84% Mismatches: 59
DB: 3 Gaps: 28
11

US-10-627-571-2 (1-188) x US-09-134-001C-1541 (1-2637)

Qy 4 AspValPheAsnSer---LysAsnLeuAlaValGlnIaGlnIleuGlyLys 22
      :::::|||||:::|||||:::
Db 1525 GAGATTTTGGTACTACTATTAAAAATATTATACAGACATCATATATAAAATTA-----AAA 1578
      :::|||||:::|||||:::

Qy 23 MetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAsp 42
      :::|||||:::|||||:::
Db 1579 GAAATTA---AAATCTAATGTTTACTGACCTAAAGATCTACACAAACGAATAA---GAA 1632
      :::|||||:::|||||:::

Qy 43 GluLeuValArgValIleThrArgGluTyThrGlnAsnLysGluAlaGluValSerIle 62
      :::|||||:::|||||:::
Db 1633 AAAATTTTAAATGTTTGGAAAGAA-----AAGAAAAAAATATTGAGGAAGAAATC 1683
      :::|||||:::|||||:::

Qy 63 LysAsnLeuIleLysThrVal-----IleLysLeuAlaIleLeuTyr 76
      :::|||||:::|||||:::
Db 1684 AATAGAGCTTTTAAATCCCTAAGATGATATAGAAAGTAAACTAAAGAAATATGACAT 1743
      :::|||||:::|||||:::

Qy 77 ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysVal 96
      :::|||||:::|||||:::
Db 1744 GAATACACTGAG---ACTCAAAAACAAATAACTTCTATGAAACCATTAAGAAACAACTTA 1800
      :::|||||:::|||||:::

Qy 97 HisGlnLeuAlaMetThrValIleSerPhe-----HisGln 108
      :::|||||:::|||||:::
Db 1801 TCTCGAGTTAAACATCTATTGAAACCTTGAAGAAATGAAAGATCAACTTAAAGAAAT 1860
      :::|||||:::|||||:::

Qy 109 ValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys---ArgGlu 127
      :::|||||:::|||||:::
Db 1861 TTAAAGAAATATTTGATGAACAACTA-----AAAATCTTAAACAGATGGGTAAAAAA 1914
      :::|||||:::|||||:::

Qy 128 MetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys-----SerHisGlyArgVal 145
      :::|||||:::|||||:::
Db 1915 ATAAATATATGAAATTTTAAAAAAAACAGTAACATTAAATTCACACCATATGCAAAATGTT 1974
      :::|||||:::|||||:::

Qy 146 AsnAsnValPheAspHisPheSerAsp 154
      :::|||||:::|||||:::
Db 1975 AATTAATTAATAGAGTTCTCTCAAGAG 2001
      :::|||||:::|||||:::

US-09-134-001C-1541
US-09-404-879A-84/C
; Sequence 84, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitchem, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FaestSeq For Windows Version 3.0
; SEQ ID NO 84
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-84

```

Alignment Scores:

Pred. No.:	0.111	Length:	571
Score:	84.50	Matches:	47
Percent Similarity:	41.05%	Conservative:	31
Best Local Similarity:	24.74%	Mismatches:	79
Query Match:	8.78%	Indels:	33
DB:	3	Gaps:	8

US-10-627-571-2 (1-188) X US-09-404-879A-84 (1-571)

QY	7	AsnserlysaenleuValaIglInalagInlysllyleuGlylsmetAlserlyS	26
Db	505	AACAAGAAAGCAAGAGATTCAACAATTCAAGAAACCTTGACAGTACTGTGACCCAG	446
QY	27	SerlleaIaThr-----LeuIleAspAspThrSerSerGluValLeuAsp	42
Db	445	CTTGACGCTTTTACTAAGACATGTCTCCCTCCAGATGATCTGACAGGGTGTATAT	386
QY	43	GluLeuTyArgValIThrArgGluTyThrGln-----AsnlyslsGluIaGluLyS	60
Db	385	GAACCTAAGAAATGGGAGAGAAAGTTAGATGATGATTCAAAGCAAGAAAGAAATT	326
QY	61	LysIleLysAsnleuIleuTystrValIlelyleuLleIleuTyArgAsnAsnGln	80
Db	325	AGACTCAAGAGATTAATTGCAGGTTCCTTAAGCAATCACTT-----AGACAGATGCC	272
QY	81	PheAsnGlnAspGluLeuAlaLeu-----MetGlnlyslPheLyslslyVal	96
Db	271	ATCCATATGGAAGATTTAAAGATTACATTTCAGGCTTGAAACATGACAGACAGATTGG	212
QY	97	HisGlnLeuAlaMetThrValValSerPheHisGlnIValAspTyThrPheAspArgAsn	110
Db	211	GAGTCCAAAGGCCGAGACAGAGTCCAGCTTCACAG-----AAG	172
QY	117	ValLeuSerArgLeuLeuAsnGluCyArgGluMetLeuHisGlnIleleGln---Arg	133
Db	172	GTCGTGATACTCTTACAGGGGAAACCAAGAACTTTTCCAGGTCAAAAGACAGCC	112
QY	136	HisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCys	155
Db	112	CACCTATATC-----CACGTTTCAAGAT	89
QY	156	GluPheLeuAlaAlaLeuTyArgAsnProPheGluAsnPheLyProHisleuGlnlyslLeu	175
Db	88	GAA---TTAGCTAAGTTGGAATCAGAACTTAAGAAGCTCAAGAACACAGTTGACTGATTTA	32
QY	176	CysAspGlyIleAsnlyslMetLeuAspGlu	185
Db	31	AGTAACTCTTTAGAAATATGTAAGCAACA	2

RESULT 44

```

US-09-338-933-84/c
/ Sequence 84, Application US/09338933
/ Patent No. 648891
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer Lynn
/ APPLICANT: King, Gordon E.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
/ TITLE OF INVENTION: OVARIAN CANCER
/ FILE REFERENCE: 210121.462C1
/ CURRENT APPLICATION NUMBER: US/09/338,933
/ CURRENT FILING DATE: 1999-06-23
/ NUMBER OF SEQ ID NOS: 312
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 84
/
/ LENGTH: 571
/
/ TYPE: DNA
/
/ ORGANISM: Homo sapien
US-09-338-933-84

```

Alignment Scores:

Pred. No.:	0.111	Length:	571
Score:	84.50	Matches:	47

Percent Similarity:	41.05%	Conservative:	3
Best Local Similarity:	24.74%	Mismatches:	7
Query Match:	8.78%	Indels:	3

US-10-627-571-2 (1-188) x US-09-338-933-84 (1-571)

[illegible]

RESULT 45

```

US-09-215-681-84/c
/ Sequence 84, Application US/09215681A
/ Patent No. 6528253
/
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Prudakis, Tony N.
/ APPLICANT: King, Gordon E.
/
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
/
/ TITLE OF INVENTION: OF OVARIAN CANCER
/
/ FILE REFERENCE: 210121.463
/
/ CURRENT APPLICATION NUMBER: US/09/215,681A
/
/ CURRENT FILING DATE: 1998-12-17
/
/ NUMBER OF SEQ ID NOS: 310
/
/ SOFTWARE: FastSeq for Windows Version 3.0
/
/ SEQ ID NO 84
/
/ LENGTH: 571
/
/ TYPE: DNA
/
/ ORGANISM: Homo sapien
US-09-215-681-84

```

Alignment Scores:

Pred. No.:	0.111	Length:	57
Score:	84.50	Matches:	47
Percent Similarity:	41.05%	Conservative:	31
Best Local Similarity:	21.74%	Mismatches:	79

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 17:21:24 ; Search time 502 Seconds

(without alignments)

2216.955 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962

Sequence: 1 MATDVNSKNIAVQAOKIL.....KPHLOKLCDCGINKMLDENI 188

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 0%
Listing first 100 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O/cgnt2.1/USPTO.spool.h/US10627571/tunat.27072005.154719.27497/app.query.fasta_1.327
-DB=N Geneseg -OPMT=fastcap -SUPFIX=ring -MINMATCH=0.1 -LOOEXT=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=LOCAL
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10627571 @CGN 1.1 470 @tunat.27072005.154719.27497 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

N_Geneseg_16Dec04:*
1: geneseg1980s:*
2: geneseg1990s:*
3: geneseg2000s:*
4: geneseg2001as:*
5: geneseg2001bs:*
6: geneseg2002as:*
7: geneseg2002bs:*
8: geneseg2003as:*
9: geneseg2003bs:*
10: geneseg2003cs:*
11: geneseg2003ds:*
12: geneseg2004as:*
13: geneseg2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	954	99.2	1892	8	ACB81105	AccB81105 Human TNF	
2	954	99.2	1915	6	ABQ79541	Abq79541 Human SCC	
3	954	99.2	1921	4	AAH14888	Aah14888 Human CDN	
4	950	98.8	1729	4	AAH14793	Aah14793 Human CDN	
5	950	98.8	1814	13	ADR14214	Adr14214 Human NF-	

6	950	98.8	2003	13	ADR14216	Adr14216 Human NF-	
7	950	98.8	2034	11	ACN44179	Acn44179 Human mRN	
8	950	98.8	2081	11	ADR14038	Adr14038 Human NF-	
9	950	98.8	58723	11	ACN44178	Acn44178 Human gen	
10	948	98.5	1943	2	AAx98013	Aax98013 Human gen	
11	948	98.5	1943	9	ADAI1570	Adai1570 Human sec	
12	948	98.5	1943	10	ADA56241	Ada56241 Gene enco	
13	920	95.6	2087	11	ACN44177	Acn44177 Mouse mRN	
14	920	95.6	62231	11	ACN44176	Acn44176 Mouse gen	
15	916	95.2	816	4	AAH07403	Aah07403 Human CDN	
16	914	95.0	544	12	ACH89719	Ach89719 Human gen	
17	651	67.7	1986	12	ADG63108	Adg63108 Novel hum	
18	640	66.5	969	10	ADC30769	Adc30769 Human nov	
19	610	63.4	527	12	ACH92216	Ach92216 Human gen	
20	593	61.6	619	6	ABs78742	AbS78742 DNA encod	
21	593	61.6	645	6	ABs78741	AbS78741 DNA encod	
22	593	61.6	2108	6	AAL39626	Aal39626 Human sec	
23	593	61.6	2186	12	ADG85659	Adg85659 Human tum	
24	593	61.6	2186	12	ADG86751	Adg86751 Human tum	
25	593	61.6	2186	13	ADG84271	Adg84271 Human tum	
26	587	61.0	587	4	AAH07641	Aah07641 Human CDN	
27	527	54.8	1113	6	ABK34943	Abk34943 Human LP2	
28	527	54.8	1165	6	AAO38699	Aao38699 Human gen	
29	527	54.8	1175	6	ABZ11655	Abz11655 Human pol	
30	527	54.8	1175	12	ADM44173	Adm44173 Novel hum	
31	527	54.8	1268	4	AAE59594	Aae59594 Human cel	
32	401	41.7	829	4	ABL07183	Ab107183 Drosophi1	
33	383.5	39.9	4406	9	ACH48309	Ach48309 Human lun	
34	324	33.7	422	9	ACH76016	Ach76016 Human gen	
35	304	31.6	500	12	AAH04721	Aah04721 Human CDN	
36	292	30.4	714	4	AAH18715	Aah18715 Human CDN	
37	292	30.4	1602	4	AAK56622	Aak56622 Human imm	
38	281	29.2	441	4	AAK56622	Aak56622 Human imm	
39	250	26.0	340	3	AAK10352	Aak10352 Human sec	
40	206	21.4	400	4	AAI92435	Aai92435 Human pol	
41	166	17.3	252	3	AAK10359	Aak10359 Human sec	
42	148	15.4	543	12	ACH78513	Ach78513 Human gen	
43	106	11.0	1339	12	ADP04796	Adp04796 Sea squit	
44	104	10.8	2943	3	AAA95803	Aaa95803 Tomato E-	
45	96.5	10.0	1694	3	AAA95805	Aaa95805 Tomato E-	
46	96.5	10.0	1694	3	AAA95812	Aaa95812 Tomato E-	
47	96.5	10.0	2154	3	ABK79037	Abk79037 Bacillus	
48	95	9.9	638	6	AAW70141	Aaw70141 Plasmodiu	
49	93.5	9.7	1809	9	ACA22957	AcA22957 Prokaryot	
50	93.5	9.7	2343	8	ACA22957	AcA22957 Prokaryot	
51	93.5	9.7	110000	2	AAx20248_08	Continuation (9 of	
52	92.5	9.6	2501	11	ACN92808	Continuation (9 of	
53	92	9.6	265118	5	AAH41227	Aah41227 Pyrococcu	
54	91.5	9.5	110000	2	AAx20248_06	Continuation (7 of	
55	91.5	9.5	111309	2	AAx20250	Aax20250 Borrelia	
56	90.5	9.4	848	4	AAH03594	Aah03594 Human CDN	
57	90.5	9.4	1359	8	ACA47387	Aca47387 Prokaryot	
58	90.5	9.4	1628	4	AAH13782	Aah13782 Human CDN	
59	90.5	9.4	2222	4	AAH75522	Aah75522 Human myo	
60	90.5	9.4	2322	2	AAV83134	Aav83134 Polynucle	
61	90.5	9.4	4256	5	ADL63542	Adl63542 Human ova	
62	90.5	9.4	4428	11	ADN39127	Adn39127 Cancer/an	
63	90.5	9.4	4500	10	ADN15655	Adn15655 Human str	
64	90.5	9.4	4501	13	ACN39836	Acn39836 Tumour-as	
65	90.5	9.4	5923	12	ADL13178	Adl13178 Human ste	
66	90.5	9.4	6633	8	ABX34762	Abx34762 Human mdd	
67	90.5	9.4	110000	10	ADF77343_00	Adf77343 Lactic ac	
68	90	9.4	1887	2	AAx99539	Aax99539 Nucleic ac	
69	90	9.4	110000	2	AAx21209_02	Continuation (3 of	
70	89.5	9.3	1140	2	AAx56021	Aax56021 Merozoite	
71	89.5	9.3	1140	2	AAx25593	Aax25593 Merozoite	
72	89.5	9.3	1142	10	ABX15269	Abx15269 P. falcip	
73	89.5	9.3	1142	11	ADM86686	Adm86686 Plasmodiu	
74	89	9.3	6538	3	AACT7192	Aac7192 Human ORF	
75	88.5	9.2	1351	4	ABL24777	Ab124777 Drosophi1	
76	88.5	9.2	1545	4	ABL19371	Ab119371 Drosophi1	
77	88.5	9.2	1700	6	ABL63353	Ab163353 Breast ca	
78	88.5	9.2	1700	6	ABL63353	Ab163353 Breast ca	

79	88.5	9.2	1700	6	ABK84326	Abk84326 Human cDN
80	88.5	9.2	1700	10	AC556483	AC556483 Human sig
81	88.5	9.2	1700	11	AD131572	AD131572 Human cDN
82	88.5	9.2	1700	11	AD156279	AD156279 Human pol
83	88.5	9.2	3204	4	AAH54121	AAH54121 S. epider
84	88.5	9.2	3227	4	AAH54947	AAH54947 S. epider
85	88.5	9.2	3760	4	AAH54665	AAH54665 S. epider
86	88.5	9.2	4202	4	AB124776	AB124776 Drosophila
87	88.5	9.2	4396	4	AB119370	AB119370 Drosophila
88	88	9.1	1638	5	AAH26501	AAH26501 Human low
89	88	9.1	1793	8	AB234807	AB234807 Human low
90	88	9.1	1993	12	AD078021	AD078021 Coding se
91	88	9.1	1893	4	AAK51890	AAK51890 Human syn
92	88	9.1	2523	4	ABA09008	ABA09008 Human pol
93	88	9.1	2523	4	AAK52874	AAK52874 Human pol
94	88	9.1	3500	10	AD039173	AD039173 Novel hum
95	88	9.1	4206	13	ADR75291	ADR75291 Bovine ac
96	88	9.1	4697	2	AAV32839	AAV32839 Human low
97	88	9.1	4697	5	AAH26496	AAH26496 Human low
98	88	9.1	4730	13	ADR75289	ADR75289 Bovine dlt
99	88	9.1	5085	8	ACC46153	ACC46153 Human dlt
100	88	9.1	6577	8	ACC46130	ACC46130 Human dlt

ALIGNMENTS

```
RESULT 1
ACC81105 1 ACC81105 standard; mRNA; 1892 BP.
AC      ACC81105;
XX
XX      25-JUL-2003 (first entry)
DE      Human TNF-induced protein GG2-1 mRNA.
XX
XX      Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; aschma;
KW      immunosuppressive; antiallergic; antiallergic; antinflammatory;
KW      lymphocyte activation; lymphocyte migration; cytokine production;
KW      cell surface marker expression; antibody production; apoptosis; allergy;
KW      antibody proliferation; antibody differentiation; hypersensitivity;
KW      graft versus host disease; inflammation; gene; ss; TNF-induced protein;
KW      GG2-1.
XX
XX      Homo sapiens.
OS
XX      WO2003029277-A2.
XX
XX      10-APR-2003.
XX
XX      02-OCT-2002; 2002WO-US031618.
XX
XX      03-OCT-2001; 2001US-0327212P.
XX
XX      (RIGE-) RIGEL PHARM INC.
XX
XX      Chu P, Li C, Liao XC, Maesuda E, Pardo J, Zhao H;
XX      WPI; 2003-363276/34.
XX      P-PSDB; ABR59710.
XX
XX      Identifying a compound that modulates T lymphocyte activation, useful for
PT      monitoring changes in cell surface marker expression, comprises
PT      contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with
PT      a compound.
XX
XX      Disclosure; Page 84; 126pp; English.
XX
XX      The invention relates to a novel method for identifying a compound that
CC      modulates T lymphocyte activation. The method comprises contacting a T
CC      cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,
CC      where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic
CC      acid that hybridises to a nucleic acid encoding a polypeptide having a
```

```
CC      sequence selected from two 606-amino acid sequence and a 415-amino acid
CC      sequence given in the specification. The method of the invention has
CC      immunosuppressive, antiallergic, antiallergic, and antinflammatory
CC      activity. The method is useful for identifying compounds that modulate
CC      lymphocyte activation, migration, and for monitoring changes in cell
CC      surface marker expression, cytokine production, antibody production,
CC      proliferation and differentiation, and apoptosis, using either cell lines
CC      or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as
CC      drug targets for compounds that suppress or activate lymphocyte
CC      activation and migration, e.g. for the treatment of diseases in which
CC      modulation of the immune response is desired such as delayed type
CC      hypersensitivity reactions, asthma, allergies, graft versus host disease,
CC      and acute and chronic inflammation. Modulators of lymphocyte activation
CC      are useful for treating disorders related T and B cell activation and
CC      migration. The present sequence is used in the exemplification of the
CC      invention
XX
XX      SQ      Sequence 1892 BP; 618 A; 318 C; 381 G; 575 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.:      4,146-98      Length:      1892
XX      Score:      954.00      Matches:      187
XX      Percent Similarity:      99.47%      Conservative:      0
XX      Best Local Similarity:      99.47%      Mismatches:      1
XX      Query Match:      99.17%      Indels:      0
XX      DB:      8      Gaps:      0
```

US-10-627-571-2 (1-188) x ACC81105 (1-1892)

```
QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaGlnAlaGlnLysLysLeu 20
DB      98 ATGGCCACAGATGCTTTATTCCAAAACCTGGCGCTTACGGCAAAAGAGATCTTG 157
QY      21 GlyLysMetValSerLysSerLysLeaThrThrLeuLysAspThrSerSerGluVal 40
DB      158 GGTAAATGCTGCTCCAAATCCATCCGCCACCTTAATAGACACAAAGTGTGGTGG 217
QY      41 LeuAspGluLeuThrArgValThrArgGluThrGlnAsnLysLysGlnLys 60
DB      218 CTGGAGAGCTTACAGGTGACCGAGGAGTACCCCAACAAAGAGGAGGAGAGAG 277
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
DB      278 ATCATCAAGAACCTCATCAAGCACTCATCAAGCTGGCATCTTTATAGATATATACG 337
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB      338 TTTAATCAAGATGATGACTGATGATGAGAAATTTAAGAGAAAGTTTCATCAGCTTGC 397
QY      101 MetThrValIleSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
DB      398 ATGACCGGTGATGATTCATCAGGTGATGATGATGATGATGATGATGATGATGATG 457
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB      458 CTGTTAATGATGATGAGAGATGCTGCACCAATCATTCAGCGCACTCATCCTGCAG 517
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerArgCysGluPheLeuAlaLys 160
DB      518 TCACATGACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 577
QY      161 LeuTyrAsnProPheGluAsnPheLysProHisGlnIleLysLeuCysAspGlyIleAsn 180
DB      578 TTGTATTAATCTTTTGGGATTTTAAACCCCATTAACAAACATATGATGATATAC 637
QY      181 LysMetLeuAspGluLysAsnIle 188
DB      638 AAAATGTTGATGAAGAGACATTA 661
RESULT 2
ABQ79541
ID      ABQ79541 standard; cDNA; 1915 BP.
XX
```

AC ABQ79541;
 XX 25-NOV-2002 (first entry)
 XX
 XX Human SCC-S2 protein encoding cDNA.
 DE
 XX SSC-S2; apoptosis; tumour; cancer; cytostatic; antisense therapy; human;
 KM gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 134..700
 FT /*tag= a
 FT /product= "SCC-S2"
 XX
 PN W0200259337-A1.
 PD 01-AUG-2002.
 PD
 PD 28-JAN-2002; 2002MO-US002212.
 PF
 XX 26-JAN-2001; 2001US-0264062P.
 PR
 XX (GEOU) UNIV GEORGETOWN SCHOOL MEDICINE.
 PA
 XX Kasid UN, Kumar D, Gokhale P, Ahmad I;
 XX WPI; 2002-657478/70.
 DR P-PSDB; ABB81967.
 XX
 XX New SCC-S2 polypeptides and nucleic acids encoding them, useful as a
 PT target for identifying compounds that modulate cancer progression by
 PT inhibiting apoptosis, as a target for detecting cancers, or for
 PT immunizing animals.
 XX
 PS Claim 2; Fig 1; 70pp; English.
 XX
 CC The invention relates to an anti-apoptotic gene SSC-S2 and encoded
 CC protein. The gene is a positive mediator of tumour growth and metastasis
 CC in certain cancer types. The SSC-S2 protein can be expressed by standard
 CC recombinant methodology. The SSC-S2 polypeptide is useful as a target for
 CC identifying compounds that modulate cancer progression by inhibiting
 CC apoptosis, as a target for detecting cancers where this polypeptide is
 CC overexpressed, e.g. renal and ovarian cancers, and leukemia. The antibody
 CC and antisense oligonucleotide can be used to treat cancer and to inhibit
 CC cancer cell proliferation and/or metastasis. The present sequence
 CC represents a human SSC-S2 protein encoding cDNA
 CC
 SO Sequence 1915 BP; 610 A; 331 C; 392 G; 582 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 4.21e-98 Length: 1915
 Score: 954.00 Matches: 187
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 99.17% Indels: 0
 DB: Gaps: 0
 US-10-627-571-2 (1-188) x ABQ79541 (1-1915)
 QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
 DB 134 ATGGCCACAGATGCTTTTATTCCAAAACCTGCGCTTCAGGCACAAAGAGATCTTG 193
 QY 21 GlyLysMetValSerLysSerIleAlaThrThrIleuLeuAspThrSerSerGluVal 40
 DB 194 GGTAAATGCTGCCAAATCATCGCCACCACTTAAATACACACAAATAGTAGG 253
 QY 41 LeuAspGluLeuValThrArgGluThrGlnAsnLysLysGlnAlaGlnLys 60
 DB 254 CTGATGAGCTCTACAGAGTGCACAGGAGTACACCAAAACAAGAAAGAGGAGAG 313

QY 61 LysIleLysAsnLeuIleuLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
 DB 314 ATCATCAGAACTCTATCAAGACAGTCATCAAGCTGCGCTTTTATAGAAATATCAG 373
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 DB 374 TTTAAATCAAGATGAGCTAGCATTCATGAGCAAAATTTAAGAAAGAAAGTTCAGCTTCT 433
 QY 101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValIleSerArg 120
 DB 434 ATGACCGTGGTTCAGTTCCATCATCGTGGATTATACCTTCACCGAATGTGTTATCCAG 493
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 DB 494 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTTAGCGCACCTCAGCTGCAAG 553
 QY 141 SerHisGluYArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
 DB 554 TCACATGAGACGGTAAATATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 613
 QY 161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 DB 614 TTGTATATATCTTTTGGGAATTTTAAACCCACTTACAAAACTATGTGATGATCAAC 673
 QY 181 LysMetLeuAspGluGluLeuIle 188
 DB 674 AAATGTGATGAGAGAAACATA 697
 DB
 RESULT 3
 ID AAH14888 standard; cDNA; 1921 BP.
 XX AAH14888;
 AC
 XX 26-JUN-2001 (first entry)
 DT
 XX Human cDNA sequence SEQ ID NO:12751.
 DE
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KM
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 PN
 XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000BP-00116126.
 PF
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Iehil S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 DR
 XX
 XX Primer sets for synthesising polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12751; 2537PP + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the


```

Db      101 GTGGCCAGACGATGCTTAATTCAAAACCTGGCCGTCAGGACAGAAAAGATCTTG 160
      21 GLLYMeTValSerlySerlleaIaThThleuIleAspAspThrSerSerGluVal 40
      161 GGTAAATGTGTGTCAAATCCATCGCCACCACTTAATTAAGCAGCAAGATGAGGTG 220
      41 LeuAspGluLeuValThrArgGluThrThGlnAsnlyslValagluys 60
      221 CTGGATGAGCTCTTACAGAGTGACCGAGGAGTACCCAAAACAGAGGAGGAGAG 280
      61 LyslleYAsnleuIleYsThrVallleYsleuIalleuYrYrAsnAsnGln 80
      281 ATCATCAAGAACTCATCAAGACAGATCATCAAGTGGCCATCTTTATAGAAATACAG 340
      81 PhAsnGlnAspGluLeuValaleuMetGluLysPheLysLysValHISGlnLeuVal 100
      341 TTTAAATCAAGATGATGAGCATTTGATGAGAAATTTAAGAAAGTTTCATCAGCTTGGT 400
      101 MetThrValValSerPheHISGlnValAspTYrThrPheAspArganValleuSerArg 120
      401 ATGACCGTGTGCTGATTTCCATCAGGTGATTTATACCTTTGACCGGAATGTGTATCCAGG 460
      121 LeuLeuAsnGluCysArgGluMetLeuHISGlnlleGlnArgHISleuThrAlaLys 140
      461 CTGTTAAATGAATGACAGAGAGATGCTGCACCAATCATTTAGCGCCACTCATGCGCAG 520
      141 SerHISGlyArgValAsnAsnValPheAspHISPheSerAspCysGluPheLeuAlaVal 160
      521 TCACATGACGAGCGGTATTAATGTTGTTGATCATTTTCAATTTGTGAATTTTGGCTGCC 580
      161 LeuTYrAsnProPheGlyAsnPheLysPheHISleuGlnLysleuCysAspGlylleAsn 180
      581 TGTATAATCTCTTTGGGAATTTTAAACCCCACTTCAAAAACATATGTATGATATCAAC 640
      181 LysMetLeuAspGluGluAsnIle 188
      641 AAAATGTTGATGAAGAGAAACATA 664

Db
RESULT 5
ADRI4214
ID      ADRI4214 standard; DNA, 1814 BP.
XX
XX      ADRI4214;
XX
XX      21-OCT-2004 (first entry)
XX
XX      Human NF-kappaB pathway-associated gene SegID215.
XX
XX      NF-kappaB pathway; antiinflammatory; cytoskeletal; hepatotropic; virucide;
XX      antiarthritic; antirheumatic; gastrointestinal-Gen; antiaesthetic;
XX      antidiabetic; immunomodulator; cerebroprotective; vasotrophic;
XX      immunosuppressive; vulnervary; gene therapy; immune disorder;
XX      inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
XX      hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
XX      hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
XX      X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
XX      viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
XX      rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
XX      atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
XX      autoimmune disorder; hyper immune activity;
XX      aberrant acute phase response; hypercongenital condition; birth defect;
XX      necrotic lesion; wound; organ transplant rejection;
XX      aberrant signal transduction; proliferating disorder; cancer;
XX      HIV propagation; gene; de; human.
XX
XX      Homo sapiens.
XX
XX      MO2004065577-A2.
XX
XX      05-AUG-2004.
XX

```

```

PF      13 -JAN-2004; 2004MO-US0000798.
XX
XX      14 -JAN-2003; 2003US-0440068P.
PR      12 -MAY-2003; 2003US-0469757P.
XX
XX      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX      Nedler SG, Neubauer MG, Feder JN, Carmen J;
XX      WPI; 2004-562168/54.
DR      P-PSDB; ADRI4215.
XX
XX      New isolated polynucleotides and polypeptides associated with NF-kappaB
PT      pathway, useful for diagnosing, treating, or preventing disorders or
PT      diseases associated with NF-kappaB pathway.
XX
XX      Claim 1; SEQ ID NO 215; 237bp; English.
XX
CC      This invention relates to the novel association of protein sequences (and
CC      the genes which encode them) to the NF-kappaB pathway. The invention may
CC      be useful for the production of compounds with an antiinflammatory,
CC      cytoskeletal, hepatotropic, virucide, antiaesthetic, antirheumatic,
CC      gastrointestinal-Gen, antiaesthetic, antidiabetic, antirheumatic,
CC      immunomodulator, cerebroprotective, vasotrophic, immunosuppressive or
CC      vulnervary activity or for gene therapy. The proteins and nucleotides are
CC      useful for diagnosing, preventing, treating, or ameliorating conditions
CC      or diseases associated with the NF-kappaB pathway. The condition is an
CC      immune disorder, an inflammatory disorder, an inflammatory disorder
CC      related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC      hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM
CC      syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC      ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC      hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC      survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC      bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC      syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC      immune activity, disorders related to aberrant acute phase responses,
CC      hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC      organ transplant rejection, conditions related to organ transplant
CC      rejection, disorders related to aberrant signal transduction,
CC      proliferating disorders, cancers and HIV propagation in cells infected
CC      with other viruses. The present sequence is that of a human gene which is
CC      subject to the novel association with the NF-kappaB pathway of the
CC      invention. Note: This sequence does not appear in the specification but
CC      was obtained by the indexer from Genbank.
XX
XX      Sequence 1814 BP; 590 A; 297 C; 372 G; 555 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 1,11e-97 Length: 1814
XX      Score: 950.00 Matches: 186
XX      Percent Similarity: 99.47% Conservative: 1
XX      Best Local Similarity: 98.94% Mismatches: 1
XX      Query Match: 98.75% Indels: 0
XX      DB: 13 Gaps: 0
XX
XX      US-10-627-571-2 (1-188) x ADRI4214 (1-1814)
XX
XX      1 MetAlaThAspValPheAsnSerlyAsnleuAlaValGlnAlaGlnlyslleu 20
XX      174 GTGGCCAGACGATGCTTAATTCAAAACCTGGCCGTCAGGACAGAAAAGATCTTG 223
      21 GLLYMeTValSerlySerlleaIaThThleuIleAspAspThrSerSerGluVal 40
      234 GGTAAATGTGTGTCAAATCCATCGCCACCACTTAATTAAGCAGCAAGATGAGGTG 293
      41 LeuAspGluLeuValThrArgGluThrThGlnAsnlyslValagluys 60
      294 CTGGATGAGCTCTTACAGAGTGACCGAGGAGTACCCAAAACAGAGGAGGAGAG 353
      61 LyslleYAsnleuIleYsThrVallleYsleuIalleuYrYrAsnAsnGln 80
      354 ATCATCAAGAACTCATCAAGACAGATCATCAAGTGGCCATCTTTATAGAAATACAG 413

```

QY 81 PheAeNgInAsPGLuLeuAlaLeuMetGLuLyPheLyLyValHisGLInLeuAla 100
Db 414 TTTAAATCAAGATGAGCTGACATTGATGAGAAATTTAAAGAAAGTTGATCAGCTTGCT 473
QY 101 MetThrValValSerPheHisGLInValAspTYrThrPheAspArgAsnValLeuSerArg 120
Db 474 ATGACCGTGGTCAAGTTTCCATCAGGTGATTAATCCTTTGACCGGAATGTGTTATCCAGG 533
QY 121 LeuLeuAsnGLuCyArgGLuMetLeuHisGLInIleIleGLInArgHisLeuThrAlaLys 140
Db 534 CTGTTAAATGAATGACAGAGATGCTGCACCAAAATCATTCAGGCGCACCTCACTGCCAAG 593
QY 141 SerHisGLyArgValAsnAsnValPheAspHisPheSerAspCyGLuPheLeuAlaIle 160
Db 594 TCACATGACGGGGTTAAATATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 653
QY 161 LeuTYrAsnProPheGLyAsnPheLysProHisLeuGLInLeuCyAspGLyIleAsn 180
Db 654 TTGTAAATCCTTTTGGGAATTTTAAACCCCACTTACAAACTATGTATGTATCAAC 713
QY 181 LysMetLeuAspGLuGLuAsnIle 188
Db 714 AAAATGTTGGATGAGAGAAACATA 737
RESULT 6
ADRI4216
ID ADRI4216 standard; DNA; 2003 BP.
AC ADRI4216;
XX 21-OCT-2004 (first entry)
DT Human NF-kappaB pathway-associated gene SeqID217.
XX
DE
XX NF-kappaB pathway; antiinflammatory; cytoskeletal; hepatotropic; virucide;
KW antiarthritic; antirheumatic; gastrointestinal-gen; antiasthmatic;
KW antiatherosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnerable; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV, influenza;
KW viral replication; host cell survival; evasion of immune response;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; gene; ds; human.
XX
OS Homo sapiens.
XX
PN WO2004065577-A2.
XX
PD 05-AUG-2004.
XX
PF 13-JAN-2004; 2004WO-US000798.
XX
PR 14-JAN-2003; 2003US-0440068P.
XX
PR 12-MAY-2003; 2003US-0469575P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Nadler SG, Neubauer MG, Feder JN, Carman J;
XX
DR MPI: 2004-562168/54.
XX
DR P-PSDB; ADRI4217.
PT New isolated polynucleotides and polypeptides associated with NF-kappaB pathway, useful for diagnosing, treating, or preventing disorders or

PT diseases associated with NF-kappaB pathway.
XX
XX Claim 1; SEQ ID NO 217; 237BP; English.
XX
XX This invention relates to the novel association of protein sequences (and
CC the genes which encode them) to the NF-kappaB pathway. The invention may
CC be useful for the production of compounds with an antiinflammatory,
CC cytoskeletal, hepatotropic, virucide, antiaarthritic, antirheumatic,
CC gastrointestinal-gen, antiasthmatic, antiatherosclerotic,
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC vulnery activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappaB pathway. The condition is an
CC immune disorder, an inflammatory disorder, an inflammatory disorder
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction,
CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human gene which is
CC subject to the novel association with the NF-kappaB pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.
XX
SQ Sequence 2003 BP; 667 A; 324 C; 406 G; 606 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,276-97 Length: 2003
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 13 Gaps: 0
US-10-627-571-2 (1-188) x ADRI4216 (1-2003)
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGLInAlaGLInLysIleLeu 20
Db 204 GTGGCCACAGATGCTCTTAATTCCTCAAAACCTGGCGCTTCAGGCCCAAAAGAGATCTTG 263
QY 21 GLYsMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGLuVal 40
Db 264 GGTAAATGGGTCCAAATTCATCGCCACACCTTAATAGACACAAAGTAGTAGGTG 323
QY 41 LeuAspGLuLeuTYrArgValThrArgGLuTYrThrGlnAsnLysLysGLuAlaGLuLys 60
Db 324 CTGATGAGCTCTTACAGAGTGACAGGAGTACACCAAAACAAGAGAGCAGAGAAG 383
QY 61 LysIleLeuAsnLeuIleLysThrValIleLysLeuAlaIleLeuTYrArgAsnGln 80
Db 384 ATCATCAAGAACCTCATCAGACAGTCAATCAAGCTGCGCATCTTTATAGAAATATCAG 443
QY 81 PheAeNgInAsPGLuLeuAlaLeuMetGLuLyPheLyLyValHisGLInLeuAla 100
Db 444 TTTAAATCAAGATGAGCTTGCATTGAGAGAAATTTAAAGAAAGTTGATCAGCTTGCT 503
QY 101 MetThrValValSerPheHisGLInValAspTYrThrPheAspArgAsnValLeuSerArg 120
Db 504 ATGACCGTGGTCAAGTTTCCATCAGGTGATTAATCCTTTGACCGGAATGTGTTATCCAGG 563
QY 121 LeuLeuAsnGLuCyArgGLuMetLeuHisGLInIleIleGLInArgHisLeuThrAlaLys 140
Db 564 CTGTTAAATGAATGACAGAGATGCTGCACCAAAATCATTCAGGCGCACCTCACTGCCAAG 623
QY 141 SerHisGLyArgValAsnAsnValPheAspHisPheSerAspCyGLuPheLeuAlaIle 160

```

Db      624 TCACATGACGGGTATTAATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 683
Qy      161 LeuTYrAsnProPheGlyAsnPhelYsProHisLeuGlnYsLeuCyAspGlyIleAsn 180
Db      684 TTGTATATATCTTTTGGGAATTTTAAACCCACCTTACAAAACTATGTGATGATCAAC 743
Qy      181 LysMetLeuAspGluGluIle 188
Db      744 AAAATGTTGATGAGGAACATA 767

RESULT 7
ACN44179
ID      ACN44179 standard; cDNA; 2034 BP.
AC      ACN44179;
XX      18-NOV-2004 (first entry)
XX      Human mRNA sequence hct1950131.
XX      Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX      Homo sapiens.
XX      MO2003073826-A2.
XX      12-SEP-2003.
XX      28-FEB-2003; 2003WO-US006235.
XX      01-MAR-2002; 2002US-00087192.
XX      (SAGR-) SAGRES DISCOVERY.
XX      Morris DW;
XX      WPI; 2003-328604/31.
XX      Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX      comprises a nucleotide sequence.
XX      Claim 1; SEQ ID NO 497; Opp; English.
XX      The present invention relates to novel DNA and protein sequences which
XX      are associated with carcinomas. The sequences are useful for: (i) for
XX      screening drug candidates; (ii) for screening of bioactive agent capable
XX      of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX      a bioactive agent capable of modulating the activity of CAP; (iv) for
XX      evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX      carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX      carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bloodip;
XX      (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX      determining Carcinoma Associated (CA) gene copy number. In addition, the
XX      CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX      carcinoma including lymphoma. The present sequence is one such CA gene
XX      sequence. Note: This patent is an equivalent to basic patent
XX      US2002182586A1, for which no sequence data was published
XX      SO      Sequence 2034 BP; 639 A; 352 C; 412 G; 631 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1,3e-97      Length:      2034
Score:          950.00      Matches:      186
Percent Similarity: 99.47%      Conservative: 1
Best Local Similarity: 98.94%      Mismatches: 1
Query Match:    98.75%      Indels:      0
DB:            11      Gaps:        0

US-10-627-571-2 (1-188) x ACN44179 (1-2034)
Qy      1 MetAlaThraSpAlaPheAsnSerLyAsnLeuAlaValGlnAlaGlnYsIleLeu 20
:::|||||

```

```

Db      248 GTGGCCACAGATGCTTTAATTCCAAAACCTGGCCGTCAGGACAAAAGAGATCTTG 307
Qy      21 GlyLeuMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db      308 GGTAATAATGGGTGTCAAATTCATCGCCACCACTTAATAGACACAAAGATGAGAGGTG 367
Qy      41 LeuAspGluLeuTYrArgValThrArgGluTYrThrGlnAsnYsGluValGluVal 60
Db      368 CTGGATGAGCTTACAGAGTACAGACGAGGAGTACCCAAACAAAGAGGAGGAGAG 427
Qy      61 LysIleLeuAsnLeuIleYsThrValIleLeuLeuAlaIleLeuTYrArgAsnGln 80
Db      428 ATCATCAAGACCTCATCAAGACAGTCATCAAGCTGCGCATTTTATAGAAATATCAG 487
Qy      81 PheAsnGlnAspGluLeuAlaLeuMetGluYsPheLeuYsValHisGlnLeuAla 100
Db      488 TTTAATCAAGATGAGCTAGATGATGAGAAATTTAAGAGAAATTCATCAGCTTGGT 547
Qy      101 MetThrValAlaSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArg 120
Db      548 ATGACCGTGCTGATTCATCCATGAGTGATTAACCTTTGACCGGAATGTGATTCAGG 607
Qy      121 LeuLeuAsnGluCyAspArgIleuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaYs 140
Db      608 CTGTAAATGATGATGACAGAGAGTGTGACCAATATTCAGGCGCACTCACTGCGAAG 667
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCyGluPheLeuAlaAla 160
Db      668 TCACATGACGAGGGTTAATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 727
Qy      161 LeuTYrAsnProPheGlyAsnPhelYsProHisLeuGlnYsLeuCyAspGlyIleAsn 180
Db      728 TTGTATATATCTTTTGGGAATTTTAAACCCACCTTACAAAACTATGATGATGATCAAC 787
Qy      181 LysMetLeuAspGluGluIle 188
Db      788 AAAATGTTGATGAGGAACATA 811

RESULT 8
ADRI4038
ID      ADRI4038 standard; DNA; 2081 BP.
AC      ADRI4038;
XX      21-OCT-2004 (first entry)
XX      Human NF-kappaB pathway--associated gene SeqID39.
XX      NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
XX      antiallergic; antirheumatic; gastrointestinal-gen; antiaesthetic;
XX      antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
XX      immunosuppressive; vulnerary; gene therapy; immune disorder;
XX      inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
XX      hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
XX      hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
XX      X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
XX      viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
XX      viral replication; host cell survival; evasion of immune response;
XX      rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
XX      atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
XX      autoimmune disorder; hyper immune activity;
XX      aberrant acute phase response; hypercongenital condition; birth defect;
XX      necrotic lesion; wound; organ transplant rejection;
XX      aberrant signal transduction; proliferating disorder; cancer;
XX      HIV propagation; gene; ds; human.
XX      Homo sapiens.
XX      MO2004065577-A2.
XX      05-AUG-2004.
XX      13-JAN-2004; 2004WO-US000798.

```


QY 61 LysIleLysAsnLeuIleuYThrValIleLysLeuAlaIleuYTrArgAsnGln 80
 DB 300 ATCATCAAGAACTCTATCAAGACAGTCATCAAGCTTCTTTATGGAATATATCG 359
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAla 100
 DB 360 TTTAAATCAAGATGAGTATGATGAGAAATTTAAGAAAGTTTCATGAGCTTGCT 419
 QY 101 MetThrValIleSerPheHisGlnValAspTrpThrPheAspArgAsnValLeuSerArg 120
 DB 420 ATACCGGTGCTGATTTCCATGAGTGATTAATCTTTGACCGGAATGTGTATCCAG 479
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 DB 480 CTGTTAAATGAATGACGAGAGATGCTGCACCAATATCATTAAGCGCCACTCATGCGCAAG 539
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
 DB 540 TCACATGACGCGGTATATATGTGTGATCATTTTTCAGATTGTGATTTTGGCTGCC 599
 QY 161 LeuTYAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 DB 600 TTGTAATATCTTTTGGGAATTTTAAACCCACTTACAAATACTATGTATGATGATAC 659
 QY 181 LysMetLeuAspGluGluAsnIle 188
 DB 660 AAAATGTTGATGACAGAGAAACATA 683
 Db
 RESULT 11
 ADA11570
 ID ADA11570 standard; DNA; 1943 BP.
 XX
 AC ADA11570;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human cDNA encoding a novel secreted protein, SEQ ID NO 98.
 XX
 KW cancer; inflammation; immune disorder; neurological disorder;
 KW blood clotting disorder; food additive; food preservative;
 KW storage capability; fat content; nutritional component; ds; gene; human.
 XX
 OS Homo sapiens.
 XX
 PN US2003055236-A1.
 XX
 PD 20-MAR-2003.
 XX
 PF 14-MAR-2002; 2002US-00097065.
 XX
 PR 18-DEC-1997; 97US-0068006P.
 PR 18-DEC-1997; 97US-0068007P.
 PR 18-DEC-1997; 97US-0068008P.
 PR 18-DEC-1997; 97US-0068003P.
 PR 18-DEC-1997; 97US-0068054P.
 PR 18-DEC-1997; 97US-0068057P.
 PR 18-DEC-1997; 97US-0068064P.
 PR 18-DEC-1997; 97US-0070923P.
 PR 19-DEC-1997; 97US-0068169P.
 PR 19-DEC-1997; 97US-0068365P.
 PR 19-DEC-1997; 97US-0068367P.
 PR 19-DEC-1997; 97US-0068368P.
 PR 19-DEC-1997; 97US-0068369P.
 PR 17-DEC-1998; 98MO-US027059.
 PR 17-JUN-1999; 99US-00334595.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Moore PA, Ruben SM, Carter KC, Shi Y, Rosen CA, Soppet DR;
 PI Kyaw H, Wei Y, Florence KA, Duan DR, Florence C, Greene JM, Feng P;
 PI Ferrite AM, Yu G, Janat F, Ni J;
 XX
 WI MPI; 2003-567105/53.

DR P-PSDB; ADA11694.
 XX
 XX New secreted HKAR24 nucleic acid molecules and polypeptides, useful for
 PT preventing, treating, or ameliorating a medical condition, such as
 PT cancer, inflammation, immune disorders, neurological and blood clotting
 PT disorders.
 XX
 PS- Claim 1; SEQ ID NO 98; 118pp; English.
 XX
 CC The invention relates to an isolated HKAR24 nucleic acid molecule. The
 CC polypeptides, nucleic acids and antibodies are useful for diagnosing a
 CC pathological condition or a susceptibility to a pathological condition,
 CC for preventing, treating, or ameliorating a medical condition, such as
 CC cancer, inflammation and other immune disorders, neurological and blood
 CC clotting disorders. The nucleic acids are also useful for chromosome
 CC identification, radiation hybrid mapping or long-range restriction
 CC mapping. The polypeptides and antibodies are useful for providing
 CC immunohistochemistry assays. The polypeptide, polynucleotide, agonist or
 CC antagonist may also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content or other
 CC nutritional components. The present sequence represents cDNA encoding a
 CC novel human secreted protein. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov.uk/sequence.html?DocID=20030055236.
 CC
 XX
 SO Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 2,066-97 Length: 1943
 Score: 948.00 Matches: 186
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 2
 Query Match: 98.54% Indels: 0
 DB: Gaps: 0
 US-10-627-571-2 (1-188) x ADA11570 (1-1943)
 QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeu 20
 DB 120 ATGGCCACGAGTGTCTTAAATTCCAAAACCTGGCGCTTCANGCAAAAGAGATCTTG 179
 QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 DB 180 GGTAAATGTGTCTCAATCCATCCGCCACCTTATAGACGACCAAGTGTAGGTG 239
 QY 41 LeuAspGluLeuTYArgValThrArgGluTYThrGlnAsnLysGlnAlaGluLys 60
 DB 240 CTGATGAGCTCTACAGAGTGACCAAGGAGTACACCAAAACAAAGAGGACAGAGAAG 299
 QY 61 LysIleLysAsnLeuIleuYThrValIleLysLeuAlaIleuYTrArgAsnGln 80
 DB 300 ATCATCAAGAACTCTATCAAGACAGTCATCAAGCTTCTTTATGGAATATATCG 359
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAla 100
 DB 360 TTTAAATCAAGATGAGTATGATGAGAAATTTAAGAAAGTTTCATGAGCTTGCT 419
 QY 101 MetThrValIleSerPheHisGlnValAspTrpThrPheAspArgAsnValLeuSerArg 120
 DB 420 ATACCGGTGCTGATTTCCATGAGTGATTAATCTTTGACCGGAATGTGTATCCAG 479
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 DB 480 CTGTTAAATGAATGACGAGAGATGCTGCACCAATATCATTAAGCGCCACTCATGCGCAAG 539
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
 DB 540 TCACATGACGCGGTATATATGTGTGATCATTTTTCAGATTGTGATTTTGGCTGCC 599
 QY 161 LeuTYAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180

Db 600 TTGTATAATCCCTTTGGGAATTTTAAACCCCACTTACAAAACATGTGATGATAC 659
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 660 AAAATGTTGATGAGAAGAACAT 683

RESULT 12
 ADA56241
 ID ADA56241 standard; DNA; 1943 BP.
 AC ADA56241;
 XX
 XX
 DT 20-NOV-2003 (first entry)
 DE Gene encoding human secreted protein #420.
 XX
 XX
 XX Immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;
 KM cytoskeletal; cerebroprotective; neuroprotective; nootropic;
 KM cardiovascular; antiarteriosclerotic; gene therapy;
 KM human secreted protein; immune disorder; inflammation;
 KM respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
 KM inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
 KM multiple sclerosis; ischaemic brain injury; Parkinson's disease;
 KM Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 KM triple helix formation; antisense gene therapy; forensic biology; ds;
 KM gene.
 XX
 OS Homo sapiens.
 XX
 PN WC02002102994-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 19-MAR-2002; 2002WC-US008278.
 XX
 PR 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 DR WPI; 2003-167512/16.
 DR P-PDB; ADA57137.
 PT
 PT New human secreted polypeptides and polynucleotides, useful for
 PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
 PT conditions, respiratory disorders, cancers, CNS disorders, or
 PT neurodegenerative disorders.
 XX
 XX
 XX Claim 21, SEQ ID NO 430; 1754bp; English.

The invention relates to 592 new human secreted polypeptides useful for
 diagnosing, treating or preventing e.g. immune disorders, inflammatory
 conditions, respiratory disorders, cancers, CNS disorders, or
 neurodegenerative disorders, or polypeptides comprising an amino acid
 sequence at least 95% identical to the new sequences. The polypeptides,
 antibodies or antibody fragments that bind to the polypeptides, nucleic
 acids encoding the polypeptides, agonists or antagonists that binds to
 the polypeptide, are useful in preparing diagnostic or pharmaceutical
 compositions for diagnosing, treating or preventing an e.g. immune
 disorders, inflammatory conditions (e.g. inflammatory bowel disease,
 nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
 allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
 (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
 disorders (e.g. Parkinson's disease or Alzheimer's disease), and
 cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
 polynucleotides are useful for chromosome identification, chromosome
 mapping, for controlling gene expression through triple helix formation
 or antisense DNA or RNA, in gene therapy, for identifying individuals
 from minute biological samples, in forensic biology, and as hybridization
 probes. The polypeptides are useful for as molecular weight markers on

CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
 CC gels, to raise antibodies, for testing biological activities, and for
 CC treating or preventing neural disorders, immune system disorders,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal, proliferative and/or cancerous diseases. This sequence corresponds
 CC to a gene encoding one of the polypeptide of the invention. Note: The
 CC sequence data for this patent did form part of the printed specification,
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;

Alignment Scores:
 Pred. No.: 2,06e-97 Length: 1943
 Score: 948.00 Matches: 186
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 2
 Query Match: 98.54% Indels: 0
 DB: 10 Gaps: 0

US-10-627-571-2 (1-188) x ADA56241 (1-1943)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGluAspGluLysIleLeu 20
 Db 120 ATGGCCACAGATGCTTTTAATTCAAAACCTGGCCGCTTCANGCACAAAAGATCTTG 179
 QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 180 GGTAAATGGTGTGCAAAATCCATCGCCACACCTTATAGACACAAAGTATGAGTG 239
 QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGluAsnLysLysGluValGluVal 60
 Db 240 CTGGATGAGCTCTTACAGAGTGACCAAGAGTGACCAACCAACCAAGAGGAGGAGAA 299
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGln 80
 Db 300 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGCGCATTTCTTATAGAAATATCG 359
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 360 TTTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAATTCATCATCGTTGCT 419
 QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
 Db 420 ATGACCGTGAGTGTCCATCATCAGGTGAGTATACCTTTGACCGGATGTGTATCCAGG 479
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGluArgHisLeuThrAlaLys 140
 Db 480 CTGTAAATGAATGACGAGAGAGATGCTGCACAAATATTTCAGGCCACTGCTGCACAG 539
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
 Db 540 TCACATGACGAGCGGTAAATATGTTGATCATTTTCAGATTGTGAATTTTGGCTGCC 599
 QY 161 LeuTyAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 600 TTGTATAATCCCTTTGGGAATTTTAAACCCCACTTACAAAACATGTGATGATAC 659
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 660 AAAATGTTGATGAGAAGAACAT 683

RESULT 13
 ACN44177
 ID ACN44177 standard; cDNA; 2087 BP.
 AC ACN44177;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Mouse mRNA sequence MCT17722.
 XX
 XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.

```

XX OS Mus musculus.
XX PN MO2003073826-A2.
XX PD 12-SEP-2003.
XX PF 28-FEB-2003; 2003WO-US006235.
XX PR 01-MAR-2002; 2002US-00087192.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX PT WPI; 2003-328604/31.
XX PS Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX PT comprises a nucleotide sequence.
XX PS Claim 1; SEQ ID NO 494; Opp; English.
XX CC The present invention relates to novel DNA and protein sequences which
XX CC are associated with carcinomas. The sequences are useful for: (i) for
XX CC screening drug candidates; (ii) for screening of bioactive agent capable
XX CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX CC a bioactive agent capable of modulating the activity of CAP; (iv) for
XX CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX CC determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX CC carcinoma including lymphoma. The present sequence is one such CA gene
XX CC sequence. Note: This patent is an equivalent to basic patent
XX CC US2002182586A1, for which no sequence data was published
XX SQ
SQ Sequence 2087 BP; 556 A; 487 C; 508 G; 536 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,4e-94 Length: 2087
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 11 Gaps: 0

US-10-627-571-2 (1-188) x ACN44177 (1-2087)
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
DB 507 GGGCTACAGATCTCTTCAATTCCAAAAACCTGGCCGTTACGACCAAAAGAGATCTCG 566
QY 21 GLYSMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 567 GCGAAATATGATCCAAATCCATCGCCACCAACGCTGATGACACCAACGACGCGAGG 626
QY 41 LeuAspGlnLeuTyrArgValThrArgLysLysLysLysLysLysLysLysLysLys 60
DB 627 CTAGATGAGCTGTACAGGGGTGACAGAGTACACCAAGCAAGAGAGGCGGAGAGAG 686
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
DB 687 GTCATCAAGAACTCTATCAAGACGGTCAATCAAGCTGCGCTCTCCACAGAACTCAAG 746
QY 81 PheAsnGlnAspGlnLeuAlaLeuMetGluLysPheLysLysLysValIleGlnLeuAla 100
DB 747 TTCATCAAGACGAGCTGCGCTCATGAGAGAGTCAAGAGAGAGTGCACGCTTGGC 806
QY 101 MetThrValValSerPheHISGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
DB 807 ATGACGGTGTGACGCTTCCACACGATGAGTACACCTTCGACCGCAATGTGTCTCAAG 866

```

```

QY 121 LeuLeuAsnGluCysArgGluMetLeuHISGlnIleIleGlnArgHISLeuThrAlaLys 140
DB 867 CTGCTGAACGAGTATGCCAGAGAGCTCTTACACGAGATCATTTACAGCGCACCTTACCGCCAG 926
QY 141 SerHISGlyArgValAsnAsnValPheAspHISpHeserAspCysGluPheLeuAlaAla 160
DB 927 TCTCACGAGACGGGTTAATATGCTTTTGACATTTTTCAGATTGATTTTGGCTGCC 986
QY 161 LeuTyrAsnProPheGluAsnPhelyProHISLeuGlnLysLeuCysAspGlyIleAsn 180
DB 987 TTGTACAAATCCCTTGGAAAGTTTAACCTCACTTACAGAACTTGGCAGCGCATCAAC 1046
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 1047 AAAATGTTGATGAAGAGAACATA 1070

RESULT 14
ACN44176
ID ACN44176 strand; DNA; 62231 BP.
XX ACN44176;
XX AC 18-NOV-2004 (first entry)
XX DT Mouse genomic sequence MCG15520.
XX DE Mouse genomic sequence MCG15520.
XX KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX OS Mus musculus.
XX PN MO2003073826-A2.
XX PD 12-SEP-2003.
XX PF 28-FEB-2003; 2003WO-US006235.
XX PR 01-MAR-2002; 2002US-00087192.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX PT WPI; 2003-328604/31.
XX PS Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX PT comprises a nucleotide sequence.
XX PS Claim 1; SEQ ID NO 493; Opp; English.
XX CC The present invention relates to novel DNA and protein sequences which
XX CC are associated with carcinomas. The sequences are useful for: (i) for
XX CC screening drug candidates; (ii) for screening of bioactive agent capable
XX CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX CC a bioactive agent capable of modulating the activity of CAP; (iv) for
XX CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX CC determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX CC carcinoma including lymphoma. The present sequence is one such CA coding
XX CC sequence. Note: This patent is an equivalent to basic patent
XX CC US2002182586A1, for which no sequence data was published
XX SQ
SQ Sequence 62231 BP; 15482 A; 12591 C; 12729 G; 16514 T; 0 U; 4915 Other;

Alignment Scores:
Pred. No.: 2,97e-92 Length: 62231
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 11 Gaps: 0

```



```

US-10-627-571-2 (1-188) x ACN44176 (1-62231)
QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
       : : : : :
Db      50651 GTGGCTCAGATGCTTCAATTCACAAAACCTGGCCGTTGAGGACAAAGAAAGATCCTG 50710
QY      21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
       : : : : :
Db      50711 GGCAAAATGGATCAATTCATGCGACACGCTGATCGACGACACGAGGAGAGGTG 50770
QY      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLysLysGluAlaGluLys 60
       : : : : :
Db      50771 CTGATGAGCTGTACAGGGTACCAAGAGTACCCAGAAACAAAGAGAGGCGGAGAGG 50830
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGln 80
       : : : : :
Db      50831 GTCATCAAGAACCTCATCAAGACGCTCATCAAGCTGGCCGCTCCACAGGAACAATCAG 50890
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
       : : : : :
Db      50891 TTCATCAAGACGAGCTGGCGCTCATGAGAAAGTTCAAGAAAGAGTGACACGCTTGC 50950
QY      101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValIleuSerArg 120
       : : : : :
Db      50951 ATGACGCTCGTCACTCCACACGAGTACAGTACCTTCGACCGCAATGTCGTCTCAGG 51010
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
       : : : : :
Db      51011 CTGCTGAACGAGTCCGAGAGCTCTCAACAGATCATTCAGCGCCACCTTACCGGCAAG 51070
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
       : : : : :
Db      51071 TCTCAGCAGCGGGTAAATTAATGCTTTGACCAATTTCAATGTGTATTTTGGCTGCC 51130
QY      161 LeuTyArgAspProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
       : : : : :
Db      51131 TTGTACATACCTCTTTGAAAGTTTAACTTCACTTACGAAACTTTCGAGCGGATAC 51190
QY      181 LysMetLeuAspGluGluAsnIle 188
       : : : : :
Db      51191 AAAATGTTGATGAGGAACATCA 51214

RESULT 15
AAH07403
ID      AAH07403 standard; cDNA; 816 BP.
AC      AAH07403;
XX
DT      26-JUN-2001 (first entry)
XX
DE      Human cDNA clone (5'-primer) SEQ ID NO:4238.
XX
KM      Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      BP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-00116126.
XX
PR      29-JUL-1999; 99JP-00248036.
XX
PR      27-AUG-1999; 99JP-00300253.
XX
PR      11-JAN-2000; 2000JP-00118776.
XX
PR      02-MAY-2000; 2000JP-00183767.
XX
PR      09-JUN-2000; 2000JP-00241899.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Isogai T, Nishikawa T, Hayashi K, Saico K, Yamamoto J,
       Iehi S, Sugiyama T, Wakamatsu A, Negai K, Otsuki T;

```

```

XX      DR      WPI; 2001-318749/34.
XX
XX      PT      Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT      length cDNAs defined in the specification, and for the detection and/or
PT      diagnosis of the abnormality of the proteins encoded by the full-length
PT      cDNAs.
XX
PS      Claim 1; SEQ ID NO 4238; 2537bp + Sequence Listing; English.
XX
XX      The present invention describes primer sets for synthesizing 5602 full-
CC      length cDNAs defined in the specification. Where a primer set comprises:
CC      (a) an oligo-dT primer and an oligonucleotide complementary to the
CC      complementary strand of a polynucleotide which comprises one of the 5602
CC      nucleotide sequences defined in the specification; where the
CC      oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC      of an oligonucleotide comprising a sequence complementary to the
CC      complementary strand of a polynucleotide which comprises a 5'-end
CC      sequence and an oligonucleotide comprising a sequence complementary to a
CC      polynucleotide which comprises a 3'-end sequence; where the
CC      oligonucleotide comprises at least 15 nucleotides and the combination of
CC      the 5'-end sequence/3'-end sequence is selected from those defined in the
CC      specification. The primer sets can be used in antisense therapy and in
CC      gene therapy. The primers are useful for synthesizing polynucleotides,
CC      particularly full-length cDNAs. The primers are also useful for the
CC      detection and/or diagnosis of the abnormality of the proteins encoded by
CC      the full-length cDNAs. The primers allow obtaining of the full-length
CC      cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC      AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC      represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC      oligonucleotides, all of which are used in the exemplification of the
CC      present invention
XX
SQ      Sequence 816 BP; 262 A; 155 C; 180 G; 212 T; 0 U; 7 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 2,816-94 Length: 816
XX      Score: 916.00 Matches: 184
XX      Percent Similarity: 97.89% Conservative: 2
XX      Best Local Similarity: 96.84% Mismatches: 2
XX      Query Match: 95.22% Indels: 2
XX      DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x AAH07403 (1-816)
QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
       : : : : :
Db      101 GTGGCCACGATGCTTAAATTCACAAAACCTGGCCGTTGAGGACAAAGAAAGATCTTG 160
QY      21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
       : : : : :
Db      161 GGTAATATGGTGTCCAAATTCATCGCCACCACTTATAGACACCAAGTACTGAGGTG 220
QY      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLysLysGluAlaGluLys 60
       : : : : :
Db      221 CTGGATGACTTCAACAGAGTGAACAGGAGTACACCAAAACAAAGAGGAGGAGAGG 280
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGln 80
       : : : : :
Db      281 ATCATCAAGAACCTCATCAAGACGCTCATCAAGCTGGCCGCTTTCATGAAATTAACG 340
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
       : : : : :
Db      341 TTTAATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGGCT 400
QY      101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValIleuSerArg 120
       : : : : :
Db      401 ATGACCGTGTCAATTCATCAAGGAGTGAATTAACCTTTGACCGGAAATGTTATCCAGG 460
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
       : : : : :
Db      461 CTGTTAAATGATGACGAGAGATGCTGCACCAATATATTCAAGCGCACCTCACTGCGCAAG 520

```

QY 141 SerHisGlyValArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla160
| | | | |
DB 521 TCACATGACGCGGTTAAATATGCTTTGATCATTTTCAGATTGTAAATTTGGCTGCC 580
| | | | |
QY 161 LeuTyrAsnProPhGlyAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleasn 180
581 TTGTATAATACCTTTTGGGAATTTTAAACCCACTTCAAAAACCTAAGTGTATGCTATCAAC 640
| | | | |
QY 181 LysMet--LeuAspGluGluAsnIle 188
| | | | |
DB 641 AAAATTGTTGGATGAAAGAACATATA 666
| | | | |
RESULT 16
ACH89719
ID ACH89719 standard; DNA; 544 BP.
XX ACH89719;
XX
XX
XX 29-JUL-2004 (first entry)
XX
XX
DE Human genome derived single exon probe #22914.
XX
XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
XX
XX US2003194704-A1.
XX
XX 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
PI
XX
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX
PS Claim 1; SEQ ID NO 22914; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above, a
CC method of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene

CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704
XX
SQ Sequence 544 BP; 175 A; 109 C; 117 G; 143 T; 0 U; 0 Other;
| | | | |
Alignment Scores:
Pred. No.: 2,78e-94 Length: 544
Score: 914.00 Matches: 179
Percent Similarity: 99.44% Conservative: 0
Best Local Similarity: 99.44% Mismatches: 1
Query Match: 95.01% Indels: 0
DB: 12 Gaps: 0
| | | | |
US-10-627-571-2 (1-188) x ACH89719 (1-544)
QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGly 21
| | | | |
DB 3 GCCACAGATGCTCTTAAATTCANAAACCTGGCCGCTTCACGACCAAAAGAGATCTTGGGT 62
| | | | |
QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
63 AAAATGCTGTCCAAATTCATCCATCCACACCTTAATATAGACACAAAGTAGTAGGTGCTG 122
| | | | |
QY 42 AspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGlnAlaGluLysLys 61
123 GATGAGCTCTTACAGAGTGACAGGAGTACACCCAAAACAAAGAGAGGACAGAGATGC 182
| | | | |
QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGlnPhe 81
183 ATCAAGAACCTTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGGAATATCACTTT 242
| | | | |
QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
243 AATCAAGATGACCTAGCATTTGATGAGAAATTTAAGAAGAAATTCATCAGCTTGCTATG 302
| | | | |
QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
303 ACCGTGCTCAGTTTCATCAGGTGATTAATACCTTTGACCGGAATGTGTATCCAGGCTG 362
| | | | |
QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
363 TTAAATGAATGACAGAGAGTGTGACCAAAATTCATTCAGCCACTTCACCAAGTCA 422
| | | | |
QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
423 CATGACGCGGTTAAATATGATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCTTG 482
| | | | |
QY 162 TyrAsnProPhGlyAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleasnLys 181
483 TATATATCCTTTTGGGAATTTTAAACCCACTTCAAAAACCTAAGTGTATGCTATCAACAA 542
| | | | |
RESULT 17
ADQ63108
ID ADQ63108 standard; cDNA; 1986 BP.
XX ADQ63108;
XX
XX
XX 07-OCT-2004 (first entry)
XX
XX
DE Novel human cDNA sequence #269.
XX
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;

KM neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
 KM cancer.
 OS Homo sapiens.
 PN EP1440981-A2.
 XX
 PD 28-JUL-2004.
 XX
 PF 21-JAN-2004; 2004EP-00001196.
 XX
 PR 21-JAN-2003; 2003JP-00102206.
 PR 09-MAY-2003; 2003JP-00133392.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Iisogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Nagai K, Irie R;
 XX WPI; 2004-535376/52.
 DR P-PSDB; ADQ65296.
 XX
 PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX
 PS Claim 1; SEQ ID NO 269; 2449pp; English.
 XX
 CC The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.
 XX
 SQ Sequence 1986 BP; 536 A; 458 C; 497 G; 495 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 1.08e-63 Length: 1986
 Score: 651.00 Matches: 114
 Percent Similarity: 89.07% Conservative: 49
 Best Local Similarity: 62.30% Mismatches: 20
 Query Match: 67.67% Indels: 0
 DB: Gaps: 0
 US-10-627-571-2 (1-188) x ADQ63108 (1-1986)
 QY 4 AspValPheAsnSerIleAlaThrIleuAlaGlnAlaGlnIleuGlyLeuMet 23
 DB 152 GATGTTTTCAGTCAAGAGCTTGGCCCTTCAAGCCGAGAAAGATTCTGACCAATA 211
 QY 24 ValSerIleSerIleAlaThrIleuAlaGlnAlaGlnIleuGlyLeuMet 43
 DB 212 GCCAGCAAACTGGGCAACATGTTGATGATGACCCGACGCGATCTTGATGAG 271
 QY 44 LeuTyrArgValThrArgGluTyrThrGlnAsnIleuGlyLeuMet 63
 DB 272 CTCTACAAATCTCAACAAAGACACACACACACAAAGAGAGCCCAAGATCATGAA 331
 QY 64 AsnIleuIleuThrValIleuLeuAlaIleuTyrArgAsnGlnPheAsnGln 83
 DB 332 GACTTATCAAGGTGGCGATCAAAATCGGATCCTTACCGGAAACAAGTTTACGCA 391
 QY 84 AspGluLeuAlaLeuMetGluValPheIleuGlyValIleuGlnAlaMetThrVal 103
 DB 392 GAGAGCTGTTATGAGAGAGTTCCGAGAGAGCTGAACACGACCGCATGACATT 451
 QY 104 ValSerPheIleuValAspTyrThrPheAspArgAsnValIleuSerArgLeuAsn 123
 DB 452 GTACAGTTCTATAGAGTGAATACACTTGTATGAGAACTGCTCTCATCTCTGAT 511

QY 124 GluCyArgGluMetLeuHisGlnIleGlnArgHisLeuThrAlaIleuSerHisGly 143
 DB 512 GAGTGAAGGAGCTGGGCACTGATGAACCTGTCACGGGACCTGACGCCAGACCGG 571
 QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTyrAsn 163
 DB 572 CGCATCAACACCTCTTAAACCACTTCCGAGTGTGAGTCTCTCCACCTCTAAGT 631
 QY 164 ProPheGlyAsnDhelysProHisLeuGlnIleuLeuCyAspGlyIleAsnIleuMetLeu 183
 DB 632 CTGATGAGAGACTGTAGCCCACTCAAGAGATTGTGAAGAAATCAATTAAGTTCTA 691
 QY 184 AspGluGlu 186
 DB 692 GATGAGAA 700
 RESULT 18
 AD30769
 ID AD30769 standard; cDNA; 969 BP.
 AC
 XX AD30769;
 XX
 DT 18-DEC-2003 (first entry)
 DE
 XX Human novel cDNA sequence, SEQ ID NO:851.
 XX
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
 KW anticancer; osteopaths; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 15q21.2; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 DR WPI; 2003-371981/35.
 DR P-PSDB; AD31740.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 PS Claim 1; SEQ ID NO 851; 1185pp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 9% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the

CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 969 BP; 264 A; 261 C; 257 G; 187 T; 0 U; 0 Other;

XX Alignment Scores:
XX Pred. No.: 7,45e-63 Length: 969
XX Score: 640.00 Matches: 112
XX Percent Similarity: 88.46% Conservative: 49
XX Best Local Similarity: 61.54% Mismatches: 21
XX Query Match: 66.53% Indels: 0
XX DB: Gaps: 0

US-10-627-571-2 (1-188) x ADC30769 (1-969)

QY 4 AspValPheasnsertlysaenleuAlaValGlnAlaGlnlysllyleuGlylyMet 23
DB 322 GATGTTTATTAGTTCAAAGAGCTTCGCTTCAAGCCCAAGAAAGATTCTGACCAAAATA 381
QY 24 ValSerlyserlilealThrThrleuIleAspAspThrsergerGluValLeuAspGlu 43
DB 382 GCCAGCAAACTGTGGCCAAACATGTTGATTGACACCGACGAGGATCTTTGATGAG 441
QY 44 LeuTyArgValThraGluTyThrGlnAsnlyslsGlnAgluylslylely 63
DB 442 CTCTACAAAGTCAACCAAGACACACACACAAAGAGAACCCCAAGATCATTGAA 501
QY 64 AsnleuIlelyThrValIlelyleuAlaIleleuTyArgAsnAsnGlnPheAsnGln 83
DB 502 GACTTATCAAGGTGGGATCAAAATCGGATCCTTACCGGAACAAACGATTAGCCAA 561
QY 84 AspGluLeuAlaLeuMetGluTyPheTyPheTyValIleGlnleuAlaMetThrVal 103
DB 562 GAGGAGCTGTATTGTGGAGAAAGTTCGGAGAAAGCTGAACACAGCCGCTATGACATT 621
QY 104 ValSerPhehisGlnValAspTyThrPheAspArgAsnValIleuserArgLeuLeuAsn 123
DB 622 GTGAGCTTATGATGAGGGAATACACTTGCATAGGAAGCTGCTCCATCTCCGCAAT 681
QY 124 GluTyArgGluUmetleuHisGlnIleIleGlnArgHisleuThrAlaIysSerHisGly 143
DB 682 GAGTGCAGAGGACCTGGTGATGAACTGGTCACGCGGACCTGACGCCCAAGACCCACGGG 741
QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysGlnPheleuAlaAlaIleuTyAsn 163
DB 742 CGGATCAACACGCTTTTAAACACTTGGCCGATGTGAGTTCCTCTCACCCCTTATAGT 801
QY 164 ProPheGlyAsnAsnleuTyProHisleuGlnlyleuTyAspGlyIleAsnlyMetLeu 183
DB 802 CTGGATGAGAGACTGTAGGCCCAACCTCAAGAGATTTGTGAAGGAATCATTAAGTTGCTC 861
QY 184 AspGlu 185
DB 862 TCCAG 867

RESULT 19
ACH92216
ID ACH92216 standard; DNA; 527 BP.
XX
XX ACH92216;
AC
XX 29-JUL-2004 (first entry)
DT
XX
DE Human genome derived single exon probe #25411.
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
OS
XX US2003194704-A1.
PN
XX 16-OCT-2003.
PD
XX 03-APR-2002; 2002US-00029386.
PF
XX 03-APR-2002; 2002US-00029386.
PR
XX 03-APR-2002; 2002US-00029386.
XX
XX (PENNY) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
PS
XX Claim 1; SEQ ID NO 25411; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC segdata.uspto.gov/sequence.html?docID=20030194704

Alignment Scores:

Pred. No.: 8,92e-58 Length: 619
 Score: 593.00 Matches: 106
 Percent Similarity: 79.14% Conservative: 42
 Best Local Similarity: 56.68% Mismatches: 39
 Query Match: 61.64% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) X ABS78742 (1-619)

```

QY      2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGly 21
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      5 GCCATGACACCTTCTGACCAACAGAGCTGGCTCTGACGCGGACGAAGAAGCTCTGAG 64
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspPheSerSerGluValLeu 41
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      65 AAGATGGCGTCCAGGACAGTGGCGGCTGCTGTGATGATACACAGCATGAGTGGTCTG 124
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      42 AspGluLeuTyrArgValThrArgLysLysThrGlnAsnLysGluAlaGlnLysLys 61
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      125 GATGACCTGTACCGCGCACAGGAGTTACCGCGGACCGGACGAGAGGCCAGAAAGATG 184
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPhe 81
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      185 CTCACAAACCTGGTCAAGGTGGCCCTGAAAGCTGGACCTGCTGCGTGGGAGACAGCTG 244
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      245 GCGGATGAGAGCTGGCGCTGGCGGCTTCCGACACCGGCGGCGCTGCTGGCCATG 304
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      305 ACGGCGGTAGACTTCCACACGAGTGACTTCACTTGAACCGGCGGCTGCGCGCGG 364
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      365 CTGCTCGAGTGGCCGACCTGCTGCACACGAGCCGCGGTGCCCACTGACCGCCAAATGCC 424
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeu 161
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      425 CACGGCGGATCAACACCGTTCGCGCACCTAGCCGACCTGCACTTCCGCTGGCGCTC 484
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      162 TyrAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLys 181
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      485 TACGGCGCCGCGAGCCCTACCGCTCCACCTCGCAGATCTCGAAGGCCCTGGCGCG 544
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      182 MetLeuAspGluGluLysIle 188
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      545 ATGCTGACGACGAGGCGAGCTCTC 565
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 21
ABS78741
ID      ABS78741 standard; DNA; 645 BP.
XX
XX
XX      ABS78741;
AC
AC
DT      16-DEC-2002 (first entry)
DT
XX
XX
DE      DNA encoding human NOVX17A protein.
XX
XX      Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
XX      Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
XX      tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
XX      Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
XX      multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;
XX      obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
XX      infertility; inflammatory bowel disease; atherosclerosis; hypertension;
XX      scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
XX      asthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
XX      bacterial infection; parasitic infection; graft-versus-host disease;
XX      cell differentiation; cell proliferation; haematopoiesis; wound healing;
XX      angiogenesis; gene; ds.
XX

```

```

OS      Homo sapiens.
XX
XX      WO200272770-A2.
XX
XX      19-SEP-2002.
XX
XX      08-MAR-2002; 2002WO-US007283.
XX
XX      08-MAR-2001; 2001US-0274281P.
XX      09-MAR-2001; 2001US-0274849P.
XX      12-MAR-2001; 2001US-0275235P.
XX      13-MAR-2001; 2001US-0275579P.
XX      13-MAR-2001; 2001US-0275601P.
XX      14-MAR-2001; 2001US-0276009P.
XX      20-MAR-2001; 2001US-0277239P.
XX      20-MAR-2001; 2001US-0277327P.
XX      20-MAR-2001; 2001US-0277338P.
XX      21-MAR-2001; 2001US-0277791P.
XX      22-MAR-2001; 2001US-0278152P.
XX      23-MAR-2001; 2001US-0278152P.
XX      26-MAR-2001; 2001US-0278894P.
XX      27-MAR-2001; 2001US-0279036P.
XX      28-MAR-2001; 2001US-0279344P.
XX      30-MAR-2001; 2001US-0280233P.
XX      02-APR-2001; 2001US-0280802P.
XX      02-MAY-2001; 2001US-0288148P.
XX      31-MAY-2001; 2001US-0294821P.
XX      31-OCT-2001; 2001US-0335302P.
XX      04-DEC-2001; 2001US-0336375P.
XX      07-MAR-2002; 2002US-00094466.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Spytek KA, Vernet CA, Tchiernev VT, Malyanar UM, Gerlach VL,
XX      Li L, Zernusen BD, Paturajan M, Gusev VY, Kekuda R, Pena CE,
XX      Zhong M, Gangoli EA, Taupier RJ,
XX
XX      WPI: 2002-713508/77.
XX      P-PsDB; ABG97497.
XX
XX      New NOVX polypeptides and polynucleotides, useful for preventing,
XX      diagnosing or treating NOVX-associated disorders, e.g. diabetes,
XX      PT sclerosis, atherosclerosis, cancer, infections, osteoporosis or
XX      Parkinson's disease.
XX
XX      Claim 22; Page 139; 26pp; English.
XX
XX      The present invention relates to a new polypeptide (NOVX). The NOVX
XX      polypeptide, nucleic acid and antibody are useful in the manufacture of a
XX      medicament for treating a syndrome associated with a human disease,
XX      CC preferably a NOVX-associated disorder. The NOVX nucleic acids,
XX      CC polypeptides and antibodies are useful for treating, preventing or
XX      CC diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
XX      CC syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
XX      CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
XX      CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
XX      CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
XX      CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
XX      CC infertility, inflammatory bowel disease, atherosclerosis, hypertension,
XX      CC scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
XX      CC asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
XX      CC infections, or graft-versus-host disease. The nucleic acids and
XX      CC polypeptides may also be used as targets for the identification of small
XX      CC molecules that modulate or inhibit e.g. neurogenesis, cell
XX      CC differentiation, cell proliferation, haematopoiesis, wound healing and
XX      CC angiogenesis, in gene therapy, in generation of antibodies that bind
XX      CC immunospecifically to NOVX substances for use in therapeutic or
XX      CC diagnostic methods. The nucleic acids are further used as hybridisation
XX      CC probes, in chromosome mapping, tissue typing, preventive medicine, and
XX      CC pharmacogenomics. The present nucleic acid sequence encodes a human NOVX
XX      CC protein of the invention
XX
XX      Sequence 645 BP; 97 A; 228 C; 219 G; 101 T; 0 U; 0 Other;
XX

```

Alignment Scores:

Pred. No.: 9.41e-58 Length: 645
 Score: 593.00 Matches: 106
 Percent Similarity: 79.14% Conservative: 42
 Best Local Similarity: 56.68% Mismatches: 39
 Query Match: 61.64% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x ABS78741 (1-645)

```

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeuGly 21
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 23 GCCATGACACCTTCACAGCACCAGAGCTGCTCTGCAGCGCCAGAAAGACTCTTCACT 82
QY 22 LysMetValSerLysSerLleAlaThrThreuleleAspAspThrSerSerGluValLeu 41
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 83 AAGATGGCGTTCACAGCAGTGTGCGCTGCTGTGATGACACAGAGTGAAGTGTCTG 142
QY 42 AspGluLeuTyArgValThrArgLutTyThrGlnAsnLysLysGluAlaGluLysLys 61
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 143 GATAGCTGTACCGCCGACACAGGAGTTCACGCGCAGCCGACAGAGCCACAGATG 202
QY 62 IleLysAsnLeuLleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGlnPhe 81
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 203 CTCAAGAACCTGGTCAAGAGTGGCCCTTAAGCTTGAGACTGCTGCTGGGGACCACTG 262
QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 263 GCGCGTGAAGAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 322
QY 102 ThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeu 121
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 323 ACCGCCCTTCAGCTTCACACAGTGTGACTTCACTTCACCGCGCGCTGCGCGCGG 382
QY 122 LeuAsnGluCysArgLysLysLeuHisGlnIleGlnHisLysLeuThrAlaLysSer 141
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 383 CTGCTCGAAGTCCCGCAGCCTGCTCACACAGCCCTGGGTCTCCACCTGACCGCCATGTC 442
QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 443 CACGCGCGCATCAACACGCTTGGCGCACCTGACGACCTGCGACTTCCTGGCGGCTC 502
QY 162 TyraAsnProPheGlyAsnPheLysPheHisLeuGlnLysLysCysAspGlyLleAsnLys 181
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 503 TACGCGCCCGCCGACGCTTACCGCTCCACCTGCGCAGATGCGAGGCGCTGCGCGG 562
QY 182 MetLeuAspGluGluAsnIle 188
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 563 ATGCTGACACAGGCGACCTC 583

```

RESULT 22

AAL39626 standard; DNA; 2108 BP.

AAL39626;

05-SBP-2002 (first entry)

Human secreted protein DNA SEQ ID NO 70.

XX Antiartherosclerotic; cyostatic; HIV; antiallergic; antianaemic;
 XX antiaesthetic; cardiac; vasotropic; neuroprotective; nootropic; SECP;
 XX immunosuppressive; antiparkinsonian; cerebroprotective; antiinflammatory;
 XX immunosuppressive; human secreted protein; cell proliferative disorder;
 XX arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;
 XX allergy; anaemia; asthma; cardiovascular disease; developmental disorder;
 XX ischaemic heart disease; congestive heart failure; neurological disorder;
 XX renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;
 XX Parkinson's disease; epilepsy; stroke; knockin humanised animal;
 XX transgenic animal; gene therapy; gene; ds.
 XX Homo sapiens.

XX W0200238602-A2.

XX 16-MAY-2002.

XX 08-NOV-2001; 2001WO-US047420.

XX 08-NOV-2000; 2000US-0247505P.

XX 09-NOV-2000; 2000US-0248642P.

XX 16-NOV-2000; 2000US-0249824P.

XX 21-NOV-2000; 2000US-0252824P.

XX 08-DEC-2000; 2000US-0254305P.

XX 18-DEC-2000; 2000US-0256448P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yve H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Walla NK;

XX Sanjwalwa M, Thornton N, Elliott VS, Lu Y, Gietzen KJ, Burford N;

XX Ding L, Hafalia AA, Tang YT, Bandman O, Warren BA, Honchell CD;

XX Lu DM, Thangavelu K, Lee S, Xu Y, Yang J, Lai PG, Tran B;

XX Ison CH, Duggan BM, Saperstein SK;

XX WPI: 2002-519296/55.

XX P-PSDB; AAO21665.

XX Human secreted proteins and polynucleotides for diagnosing, treating or

XX preventing disorders of cell proliferative, cardiovascular,

XX developmental, neurological and autoimmune/inflammatory disorders.

XX Claim 5; Page 195-196; 229pp; English.

XX The invention relates to an isolated human secreted protein (SECP)

XX polypeptide from 63 fully defined protein sequences given in the

XX specification. The polypeptide is useful for the diagnosing/treating of a

XX disease with decreased/overexpression of SECP. Examples of disorders

XX associated with abnormal expression of SECP include a cell proliferative

XX disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory

XX disorder e.g. allergies, anaemia, asthma; cardiovascular disease e.g.

XX congestive heart failure, ischaemic heart disease; developmental disorder

XX e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.

XX Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.

XX The SECP polynucleotide and polypeptide are further useful for analysing

XX the proteome of a tissue or a cell type. The polynucleotide is useful for

XX creating knockin humanised animals (pigs) or transgenic animals (mice or

XX rats) to model human disease, and for somatic or germ-line gene therapy,

XX CC and further for generating hybridisation probes useful in mapping the

XX CC naturally occurring genomic sequence. This polynucleotide sequence

XX represents the DNA of a human secreted protein of the invention

XX Sequence 2108 BP; 408 A; 602 C; 636 G; 462 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.47e-57 Length: 2108
 Score: 593.00 Matches: 106
 Percent Similarity: 79.14% Conservative: 42
 Best Local Similarity: 56.68% Mismatches: 39
 Query Match: 61.64% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x AAL39626 (1-2108)

```

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeuGly 21
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 427 GCCATGACACCTTCACAGCACCAGAGCTGCTCTGCAGCGCCAGAAAGACTCTTCACT 486
QY 22 LysMetValSerLysSerLleAlaThrThreuleleAspAspThrSerSerGluValLeu 41
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 487 AAGATGGCGTTCACAGCAGTGTGCGCTGCTGTGATGACACAGAGTGAAGTGTCTG 546
QY 42 AspGluLeuTyArgValThrArgLutTyThrGlnAsnLysLysGluAlaGluLysLys 61
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 547 GATAGCTGTACCGCCGACACAGGAGTTCACGCGCAGCCGACAGAGCCACAGATG 606

```


QY 162 TyraSPProPhelGlyAanPhelyProHisteuGlnylsLeuCyAspGlyIleAsnIys 181
DB 593 TACGGCCCGCCGACCCCTACCGCTCCACCTGGCGAGATCTCGAGGCTCGGCCGG 652
QY 182 MetLeuAspGluGluAsnIle 188
DB 653 ATGCTGACGAGGCGACGCTC 673
RESULT 24
ADQ86751
ID ADQ86751 standard; cDNA; 2186 BP.
XX
AC ADQ86751;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #3626.
XX
KM human; tumour-associated antigenic target; TAT; cytosolic; gene therapy;
KM cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUN-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 3626; 5504pp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a) - (c); or (e) a sequence that hybridises to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a

CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytosolic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.69e-57 Length: 2186
Score: 593.00 Matches: 106
Percent Similarity: 79.14% Conservative: 42
Best Local Similarity: 76.68% Mismatches: 39
Query Match: 61.64% Indels: 0
DB: 12 Gaps: 0
US-10-627-571-2 (1-188) x ADQ86751 (1-2186)
QY 2 AlaThrAspValPheAsnSerIysAsnIleuAlaValGlnIleGlnIleIleuGly 21
DB 113 GCCATGGACACCTTCAGCACCAAGAGCTGGCTCTGAGCGGAGAAAGCTCTGAGT 172
QY 22 LysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
DB 173 AAGATGGCGCTCCAAAGCAGGAGTGGCGGCTGGTGGATGACACCAAGTGGTGGT 232
QY 42 AspGluLeuValThrArgValThrArgIleuValThrGlnAsnIysIysGlnIleuIysIys 61
DB 233 GATGACTGTACCGCCGACCAAGGAGTTCACCGGACCGGACGAGAGGCCGAGAGATG 292
QY 62 IleIysAsnIleuIleuValIleuValIleuValIleuValIleuValIleuValIleu 81
DB 293 CTCAGAAACCTGTGACAGGTGGCGCTGAAAGCTGGAGCTGTGCTGGTGGAGCCAGCTG 352
QY 82 AsnGluAspGluLeuValLeuMetGluValPheIysIleuValIleuValIleuValIleu 101
DB 353 GGGCGTGAAGAGCTGGCGCTGTGCGGCGCTTCCGACCGGCGGCGCTGCGCATG 412
QY 102 ThrValIleuSerPheIleuValIleuValIleuValIleuValIleuValIleuValIleu 121
DB 413 ACGGCGGTGACCTTCACCAAGGAGTTCACCTTCGACCGGCGGCGCTGCGCATG 472
QY 122 LeuAsnGluValThrArgIleuValIleuValIleuValIleuValIleuValIleuValIleu 141
DB 473 CTGCTGAGAGTGGCGGACCTGTGACCAAGCGGCTGGGCTCCACCTGACCGGCAAGTCC 532
QY 142 HisGluValValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleu 161
DB 533 CAGGCGCGCATCAACACGCTGTTCGCGCACCTTACCGGACCTTCCTGCTGGGCTC 592
QY 162 TyraSPProPhelGlyAanPhelyProHisteuGlnylsLeuCyAspGlyIleAsnIys 181
DB 593 TACGGCCCGCCGACCCCTACCGCTCCACCTGGCGAGATCTCGAGGCTCGGCCGG 652
QY 182 MetLeuAspGluGluAsnIle 188
DB 653 ATGCTGACGAGGCGACGCTC 673
RESULT 25
ADQ84271
ID ADQ84271 standard; cDNA; 2186 BP.
XX
AC ADQ84271;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1085.

PI Iehi S, Sugiyama T, Makamatsu A, Nagai K, Oseuki T;
 XX WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 1; SEQ ID NO 4476; 2537bp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification, where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 XX Sequence 587 BP; 178 A; 146 C; 140 G; 116 T; 0 U; 7 Other;
 SQ
 Alignment Scores:
 Pred. No.: 3,996-57 Length: 587
 Score: 587.00 Matches: 123
 Percent Similarity: 93.94% Conservative: 1
 Best Local Similarity: 93.18% Mismatches: 7
 Query Match: 61.02% Indels: 1
 DB: 4 Gaps: 0
 US-10-627-571-2 (1-188) x AAH07641 (1-587)
 QY 1 MetAlThraspValPheasSerLysAsnLeuAlaValGlnAlaGlnLysLleLeu 20
 DB 140 ATGGCCACAGATGCTTTAAATCCAAACCTGCGCTCAGGACAAAGAAATCTTG 199
 QY 21 GILYsMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 DB 200 GGTAAATGCTGTCAAATCCATCGCACCACTTAATACACACAACTATGAGGTG 259
 QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLys 60
 DB 260 CTGATGACGCTCTACAGAGTACAGGAGCTACACCAAAACAAAGAGGACGAGAG 319
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
 DB 320 ATATCAAGAACCTCATCAAGACAGATCATCAAGCTGCGCTCTTAATAGAAATATCAG 379
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 DB 380 TTTAATCAAGATGAGTACATTGATGAGCAATTTAAGAAAGAAAGTTATCAGCTTGT 439
 QY 101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 DB 440 ATGACCGTGGNCGATTCCAAAGAGGGGTTTAACCTTTGACCGAATGTGTATCCAG 459
 QY 121 LeuLeuAsnGluCys-ArgGluMetLeuHisGln 131
 DB 500 CTGTTAAATGAATGCANNAAGAAATGCTGCACCA 533

RESULT 27
 ABK34943
 ID ABK34943 standard; cDNA; 1113 BP.
 XX
 AC ABK34943;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA encoding secreted protein #81.
 XX
 KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
 KW tissue regeneration; wound healing; burn; haematopoiesis;
 KW myeloid cell deficiency; lymphoid cell deficiency.
 XX
 OS Homo sapiens.
 XX
 PN WO200177288-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US010224.
 XX
 PR 06-APR-2000; 2000US-0195582P.
 XX
 PA (GENM) GENNETICS INST INC.
 XX
 PI Wong GQ, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;
 PI WPI; 2002-179321/23.
 DR
 XX
 XX Five hundred and ninety two polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for treating
 PT immune deficiencies and disorders such as autoimmune disorders.
 XX
 PS Claim 1; Page 109; 372pp; English.
 XX
 CC The invention relates to 592 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins. The polynucleotides can be used as probes for the
 CC identification and isolation of full length cDNA and genomic DNA. The
 CC polynucleotides and proteins can also be used as nutritional supplements.
 CC The proteins are useful in the treatment of various immune deficiencies
 CC and disorders such as viral infections, bacterial infections, fungal
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention
 XX
 SQ Sequence 1113 BP; 282 A; 311 C; 304 G; 216 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,976-50 Length: 1113
 Score: 527.00 Matches: 99
 Percent Similarity: 77.22% Conservative: 40
 Best Local Similarity: 55.00% Mismatches: 39
 Query Match: 54.78% Indels: 2
 DB: 6 Gaps: 1
 US-10-627-571-2 (1-188) x ABK34943 (1-1113)

```

QY 6 PheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlySerIlyLeuGlnIlySerValSer 25
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 TTCAGTCAAAAGAGCCTGGCACTGCACAGACAGAAAGACTTGAAGTAAATGGCGGT 131
QY 26 LysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTy 45
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 132 CGCTGTGGCTCATCTCTTCATAGATGAGACAAAGCAGTAGGTGCTAGATGAGCTCTAC 191
QY 46 ArgValThrArgGluThrThrGlnAsnIlyValGlnAlaGlnIlySerIlyLeuValLeu 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 CGTGTCACAAAGAGTACAGCAGCAGCGCGCCAGCCCGCTGATCAAGAGCCTG 251
QY 66 IleValThrValIleIlyLeuAlaIleLeuTyArgAsnAsnGlnPheAsnGlnAspGlu 85
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 ATCAAAAGTGGCCATCAAGGTGGCTGTGCTGCACCGCAATGGCTTGGCCCAAGTAG 311
QY 86 LeuAlaLeuMetGlnIlyPheIlySerIlyValIleGlnLeuAlaMetThrValValSer 105
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 312 CTGGCCCTGGCTACCGGCTTTCGCCAAGAGCTGGCAGGGGTCCATGACGGCACTTAC 371
QY 106 PheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 372 TTGTGAGGTACACTTCACCTTCGAGGCTGTCTGTCTGCTGCTGACCGAGTGC 431
QY 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIlySerHisGlnArgVal 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 432 CGGAGTGTCTGCTTAAGTTGGTGGAGAACACACCTTACCGCCCAAGTCAATGGCCGATC 491
QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTyArgProPhe 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 492 CGCAGGTGTTGATCATCTCTCTGACCCAGAGTCTGTCTACCGGCCCTATAGGGCT--- 548
QY 166 GlyAsnPheIlyProHisIleuGlnIlyValLeuCysAsnGlyIleAsnIlyMetLeuAspGlu 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 549 ---GACTTCACTCAGCACCCTTGGCAAGATCTGTGACGAGCTCAGAAAGCTTAAAGCGAA 605

RESULT 28
AAD38699
ID AAD38699 standard; cDNA; 1165 BP.
XX
AC AAD38699;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human LP237 secreted protein encoding cDNA.
XX
KW Human; secreted protein; atherosclerosis; Alzheimer's disease; LP237;
KW diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
KW rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
KW reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
KW gene therapy; neoplasm; transgenic; psoriasis; ischemia; carcinoma;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
XX CDS FT 100..654
XX FT /*tag= a
XX FT /product= "Human LP237 secreted protein"
XX FT sig_peptide FT 100..180
XX FT /*tag= b
XX FT mat_peptide FT 181..651
XX FT /*tag= c
XX FT /product= "Mature human LP237 secreted protein"
XX
PN WO200226801-A2.
XX
PD 04-APR-2002.
XX
XX 14-SEP-2001; 2001WO-US026026.
XX
PR 28-SEP-2000; 2000US-0236088P.
XX

```

```

PA (ELI ) LILLY & CO ELI.
XX
XX Su EW, Wang H;
XX
XX WPI; 2002-471259/50.
XX
XX P-PSDB; AAE23984.
XX
XX
XX Novel proteins and polynucleotides of secreted proteins useful for
PT treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
PT diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury.
XX
XX
XX Claim 1; Page 140-141; 145pp; English.
XX
XX The invention relates to human secreted polypeptides designated LP095,
CC LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic acid
CC molecules encoding such polypeptides. Novel secreted proteins of the
CC invention are used for treating diseases such as atherosclerosis,
CC Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
CC arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
CC combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,
CC reperfusion injury, neoplasms and cancer especially liver cancer. They
CC are also used for wound healing. Polynucleotides of the invention can be
CC used to generate transgenic animals or knock out animals, which in turn,
CC are useful in the development and screening of therapeutically useful
CC reagents for use in the treatment of diseases associated with LP
CC polypeptide associated activity. They are also used in gene therapy. The
CC present sequence is human LP237 secreted protein encoding cDNA
XX
XX
SQ Sequence 1165 BP; 305 A; 319 C; 312 G; 229 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 6,34e-50 Length: 1165
Score: 527.00 Matches: 99
Percent Similarity: 77.22% Conservative: 40
Best Local Similarity: 55.00% Mismatches: 39
Query Match: 54.78% Indels: 2
DB: Gaps: 1
XX
US-10-627-571-2 (1-188) x AAD38699 (1-1165)
QY 6 PheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlySerIlyLeuGlnIlySerValSer 25
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 109 TTCAGTCAAAAGAGCCTGGCACTGCACAGACAGAAAGACTTGAAGTAAATGGCGGT 168
QY 26 LysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTy 45
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 CCCTGTGGCTCATCTCTTCATAGATGAGACAAAGCAGTAGGTGCTAGATGAGCTTAC 228
QY 46 ArgValThrArgGluThrThrGlnAsnIlyValGlnAlaGlnIlySerIlyLeuValLeu 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 CGTGTCACAAAGAGTACAGCAGCAGCGCGCCAGCCCGCTGATCAAGAGCCTG 288
QY 66 IleValThrValIleIlyLeuAlaIleLeuTyArgAsnAsnGlnPheAsnGlnAspGlu 85
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 289 ATCAAAAGTGGCCATCAAGGTGGCTGTGCTGCACCGCAATGGCTTGGCCCAAGTAG 348
QY 86 LeuAlaLeuMetGlnIlyPheIlySerIlyValIleGlnLeuAlaMetThrValValSer 105
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 349 CTGGCCCTGGCTACCGGCTTTCGCCAAGAGCTGGCAGGGGTCCATGACGGCACTTAC 408
QY 106 PheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 409 TTGTGAGGTACACTTCACCTTTCAGAGCTGTCTGTCTGCTGCTGCTGACCGAGTGC 468
QY 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIlySerHisGlnArgVal 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 469 CGGAGTGTCTGCTTAAGTTGGTGGAGAACACACCTTACCGCCCAAGTCAATGGCCGATC 528
QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTyArgProPhe 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 529 CGCAGGTGTTGATCATCTCTTGAAGCTTCAAGGCTCAGCGGCCCTTATAGGGCT--- 585
QY 166 GlyAsnPheIlyProHisIleuGlnIlyValLeuCysAsnGlyIleAsnIlyMetLeuAspGlu 185

```

Db 586 ---GACTTCACAGCAGCTTGGCAAGATCTGTGAACGACCTGAGAACCGAA 642

RESULT 29
ABZ11655
ID ABZ11655 standard; cDNA; 1175 BP.
XX AC ABZ11655;
XX XX
XX 20-JAN-2003 (first entry)
XX DE Human polynucleotide SEQ ID NO 537.

KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nocotropic; neuroprotective;
KW antiparkinsonian; antididiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnereary; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX XX
XX OS Homo sapiens.
XX PN WO200270539-A2.
XX PD 12-SEP-2002.
XX PE 05-MAR-2002; 2002MO-US005095.
XX PR 05-MAR-2001; 2001US-00799451.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Zhou P, Goodrich RM, Asundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX WP1: 2002-759812/82.
DR P-PSDB; ABP69438.

PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX XX
PS Claim 1; SEQ ID NO 537; 1012bp + Sequence Listing; English.

CC The invention relates to an isolated polynucleotide (1) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ112066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences

XX SQ Sequence 1175 BP; 306 A; 324 C; 313 G; 232 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6 41e-50 Length: 1175
Score: 527.00 Matches: 99
Percent Similarity: 77.22% Conservative: 40
Best Local Similarity: 55.00% Mismatches: 39
Query Match: 54.78% Indels: 2

DB:		6	Gap:	1
US-10-627-571-2 (1-188) x ABZ1655 (1-1175)				
OY	6	PheA	senSerLyAsnLeuAlaValGlnAGlnIleuGlyMetValSer	25
Dd	122	TTCACCTCAAAAGCCTGGACCTGCACAGACAGAAAGACTTACTGATGAAGATGGCGGT	181	
OY	26	LysSerIleAlaThrThrLeuIleAspAspTrpSerSeriValLeuaspGluLeuTyr	45	
Dd	182	CgcCTcTGgGCTcATcCTcTTcTAAGTAcTAGAcAGAcAGAgGTcGTAcGTAcGCTTAc	241	
OY	46	ArgValI ThrArgGlnTyrThrGlnasnIyLysGlnAlGlnIyLysIleLysAsnLeu	65	
Dd	242	cGTgTGTCCAAAGAGATAcACGACAGCCGCGCCACGCGCTGATCCAAGAcCTG	301	
OY	66	IleLysThrValIleLysLeuAlaleuTyrAdgAsnAsnGlnPheAsnGlnAspGlu	85	
Dd	302	ATCAAAGTGGCCATCAAGGTGGCTGTGCACGCCAATGGCTCTTTGGCCCCAGTAG	361	
OY	86	LeuAlaLeuMetGluLysPheLysIyLysValHisGlnLeuAlaMetThrValValSer	105	
Dd	362	CTGGccCTGGCTACCcCGcCTTTCGCCAGAAGCTGGCGGAGGGTGCATTAGAcGCACTTAc	421	
OY	106	PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuaspGluCyS	125	
Dd	422	TTTTGGTGAAGTAGACTTACCTTCAGAGGTGGTGTCTGGCTGGCCTGGCTGACCGAGTGC	481	
OY	126	ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal	145	
Dd	482	CGGGATGTGGCTGTAGAGTGGTGGAAACACACCTCACGCCCAAGTCAACATGGCCCCCATC	541	
OY	146	AsnAsnValPheAspHisPheSerAspCySGluPheLeuAlaIleuTyrAsnProPhe	165	
Dd	542	CGCCACGGTGTGATCACTTCTTACCCAGGTCTGCTACCGGCCTCTTAGTGGGCT---	598	
OY	166	GlaAsnPheLysPrcHisLeuGlnIyLysLeuCysAspGlyIleAsnIyMetLeuAspGlu	185	
Dd	599	---GACTTCACACACACACTTGGCAAGATCTGTGACGGACTCAGGAAGCTGCTAGACGA	655	
RESULT 30				
ID	ADM44173	standard, cDNA, 1175 BP.		
XX	ADMM44173;			
DT	03-JUN-2004	(first entry)		
DE	Novel human arginine-rich protein cDNA #537.			
KM	ss; gene; human; arginine-rich protein; cancer; inflammation;			
XX	genetic disorder.			
OS	Homo sapiens.			
PN	US2004053250-A1.			
PD	18-MAR-2004.			
PF	21-NOV-2002; 2002US-00302172.			
PR	05-MAR-2001; 2001US-00799451.			
PR	05-MAR-2002; 2002MO-US005095.			
PR	20-AUG-2002; 2002US-00225251.			
XX	(TANG/) TANG Y T.			
PA	(XUEA/) XUE A.			
XX	(DRMA/) DRMANNC R T.			
PI	Tang YT, Xue A, Drmanac RT;			
DR	WPI, 2004-238579/22.			


```

Db      172 CGCCTGTGCTATCTCTTCATGAGACACAGCTAGGCTGATGAGCTTAC 231
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      46 ArgValThrArgIuYrThrGlnAsnLysGluAgluYslySileYsAsnLsu 65
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      232 CGTGTGTCAAGAGTACACGACACGCCGCCGCCGCGATGATCAAGACTG 291
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      66 IleYsThrValIleYsLeuAlaIleLeuYrArgAsnAsnGlnPheAsnGlnAspGlu 85
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      292 ATCAAGGTGGCCATCAAGGTGGCTGTGTCACCGCATGCTCTTGGCCCCAGAG 351
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      86 LeuAlaLeuMetGluYsPheYsLeYsValHisGlnLeuAlaMetThrValSer 105
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      352 CTGGCCTTGCTGCTACCCCTTTCGACAGAGCTGGCGAGGGTGCATGACGAC 411
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      106 PheHisGlnValAspYrThrPheAspArgAsnValLeuSerArgLeuAsnGluCys 125
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      412 TTTCGTGAGGTAGACTTCACCTTCGAGGCTGCTTCTGGCTGGCTGACCGACTGC 471
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaYsSerHisGlyArgVal 145
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      472 CGGATGTGCTGTAGAGTTGTGTGAACACCACTCAGCCCAAGTCACATGGCGCATC 531
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuYrAsnProPhe 165
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      532 CGCCACGATGTTGATCTCTCTCTGACCCAGCTGTCTCAGGCTCTCTAGGAGCT-- 588
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      166 GlyAsnPheYsProHisLeuGlnLysLeuYsAspGlyIleAsnLysMetLeuAspGlu 185
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      589 ---GACTTCACTACGACACCTTTGGCAGATCTGTGACGACTCAGAAAGCTCTAGACGAA 645
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 32
ABLO7183
ID      ABL07183 standard; cDNA; 829 BP.
XX
AC      ABL07183;
XX
DT      26-MAR-2002 (first entry)
XX
DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 16031.
XX
KW      Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical; gene; ss.
XX
OS      Drosophila melanogaster.
XX
PN      WO200171042-A2.
XX
PD      27-SEP-2001.
XX
PF      23-MAR-2001; 2001WO-US009231.
XX
PR      23-MAR-2000; 2000US-0191637P.
XX      11-JUL-2000; 2000US-00614150.
XX
PA      (PEKE ) PE CORP NY.
XX
PI      Venter JC, Adams M, Li PWD, Myere EW,
XX
DR      WPI; 2001-656860/75.
XX
DX      P-PSDB; ABB63080.
XX
PT      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions.
XX
PS      Claim 1; SEQ ID NO 16031; 21PP + Sequence Listing; English.
XX
CC      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention

```

```

CC      discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC      sequences (ABLO1840-ABLO16175) and the encoded proteins (ABB57737-
CC      ABB72072). The sequence data for this patent did not form part of the
CC      printed specification, but was obtained in electronic format directly
CC      from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 829 BP; 229 A; 208 C; 210 G; 182 T; 0 U; 0 Other;
          Alignment Scores:
          Pred. No.:      8,09e-36      Length:      829
          Score:      401.00      Matches:      79
          Percent Similarity:      65.43%      Conservative:      44
          Best Local Similarity:      42.02%      Mismatches:      65
          Query Match:      41.68%      Indels:      0
          DB:      4      Gaps:      0
          US-10-627-571-2 (1-188) x ABL07183 (1-829)
          1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAgluYslySileLeu 20
          263 ATGGCGACCAATGTCTTCAAGTGCACAGATATGCGGCTGGCGCGAGAAAGATCTTC 322
          21 GlyLeuMetValSerLysSerIleAlaThrThrLeuIleAspAspHisSerSerGluVal 40
          323 TCAGCATGGCCCAAAAGAAATCGCAAGACCTTCATGATGCGACAGAGGCTGCTG 382
          41 LeuAspGluLeuYrArgValThrArgIuYrThrGlnAsnLysGluAgluYslySile 60
          383 CTGACACACCTCTACAGAGCTCTGCAGAGTGCACACGGGGAACAGGCGCAAGGAGAG 442
          61 LysIleYsAsnLeuIleYsThrValIleYsLeuAlaIleLeuYrArgAsnGln 80
          443 CTGATCAAGAAATCATCAAGATTGTGATCAAGATGTGTGCTCCACCGAAACAATCAG 502
          81 PheAsnGlnAspGluLeuAlaLeuMetGluYsPheYsLysValHisGlnLeuAla 100
          503 TTCAGGACGAGGAGTGCAGAAAGCGGAGCTCTTAAAGAAAGTTTCAAAACACGCGAG 562
          101 MetThrValValSerPheHisGlnValAspYrThrPheAspArgAsnValLeuSerArg 120
          563 CTGTCCATCATATCATCTTACAGAGTGCACACTTCACCTGACCTGCCGTRACCTCAAAAG 622
          121 LeuLeuAsnGluCysArgLysLeuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaYs 140
          623 TCATATAGCCGAATCGCAAGTGGCGCTGAAGTGCATTTGTGCAGCGCATCTCAGAAAG 682
          141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
          683 TCCTGTGGCCGATCGACGAGGTGTTCATTTCTTGGCGAGGAGCGCTGTGAGAGCG 742
          161 LeuYrAsnProPheGlyAsnPheYsPheHisLeuGlnLysLeuYsAspGlyIleAsn 180
          743 GCTTTCGACACCAATTCCTCCCTTACCGCGAGGTGATGGAAGATTGTGGCGACATTAAC 802
          181 LysMetLeuAspGluGluAsnIle 188
          803 GCAGCATGGAAGACGGAGATATC 826
          Db
          RESULT 33
          ABL07182/c
          ID      ABL07182 standard; cDNA; 4406 BP.
          XX
          AC      ABL07182;
          XX
          DT      26-MAR-2002 (first entry)
          XX
          DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 16028.
          XX
          KW      Drosophila; developmental biology; cell signalling; insecticide;
          XX      pharmaceutical; gene; ss.
          XX
          OS      Drosophila melanogaster.

```

PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW,
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB63079.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 16028; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at fcp.wipo.int/pub/published_pct_sequences
 XX
 SO Sequence 4406 BP; 1178 A; 1021 C; 1027 G; 1180 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7.06e-33 length: 4406
 Score: 383.50 Matches: 79
 Percent Similarity: 59.90% Conservative: 45
 Best Local Similarity: 38.16% Mismatches: 64
 Query Match: 39.86% Indels: 19
 DB: 4 Gaps: 1
 US-10-627-571-2 (1-188) x ABL07182 (1-4406)
 QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlyslleu 20
 DB 1624 ATGGCGGACAAATCTTCAAGTGCACGATATCGGGCTGGGCGGAGAAAGATCTTC 1565
 QY 21 GlyIysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 DB 1564 TCACGATGGCCACAAACAAACATCCGACAGACCTTCATGATGACGACGCGCTCGTG 1505
 QY 41 LeuAspGluLeuValThrValThrArgGluIlyThrGlnAsnIlyslGlnAlaGluIly 60
 DB 1504 CTGGACAAACCTTCAAGCTCTGACAGATGCACACGGGAGCAAGGCCAAGGGGAGAG 1445
 QY 61 IysIleIlyAsnLeuIleIlyslThrValIleIlyslLeuAlaIleuIlyThrGlnAsnGln 80
 DB 1444 CTGATCAAGAACATCTCAAGATGTGATCAAGATGGTGCTCCACCGGAAACATTCAG 1385
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluIlyPheIlyslIlysl 95
 DB 1384 TTCAGCGACGAGAGCTGCAGAAAGCGGAGCTTTTAAGAAAGTTTCACTGAGTCTT 1325
 QY 96 ValHisGlnLeuAlaMet 101
 DB 1324 AAATGTAATCTTAACATGATGATATGATTTTGTGAGAAACAGCAGCTG 1265
 QY 102 ThrValValSerPheHisGlnValAspIlyThrPheAspArgAsnValLeuSerArgLeu 121
 DB 1264 TCCATCATATTCATTCACGAAGGAGCTTCACTGACCTGCGCTACCTCGCAAAAGTCA 1205

QY 122 LeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnIleArgHisLeuThrAlaIysSer 141
 DB 1204 ATAGCCGAATGCCAAGTGGCGCTGAAGTCGATTTGTCAGCCGACATCACGAAAGTGC 1145
 QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleu 161
 DB 1144 CTGGCCCGATCGACAGAGTGTTCATTTCTTCGCGAGGAGGCCCTGTCGAGACGGCT 1085
 QY 162 TyrAsnProPheGlyAsnPheIlyProHisLeuGlnIlyslLeuCyAspGlyIleAsnIly 181
 DB 1084 TTCACACGAGATTCCTCCGCGAGGTGATGGAGAAAGATTGCGCGACATTAAGCA 1025
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 1024 GCCATGAAACGCGAGATATC 1004
 RESULT 34
 ACH48309
 ID ACH48309 standard; cDNA; 422 BP.
 XX
 AC ACH48309;
 XX
 DT 13-OCT-2003 (first entry)
 XX
 DE Human lung tumour cDNA #442.
 XX
 KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN US2003073623-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 30-JUL-2001; 2001US-00918995.
 XX
 PR 30-JUL-2001; 2001US-00918995.
 XX
 PA (DRMA/) DRMANAC R T.
 PA (LABA/) LABAT I.
 PA (STAC/) STACHE-CRAIN B.
 PA (DICK/) DICKSON M C.
 PA (JONE/) JONES L W.
 XX
 PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 DR WPI; 2003-615964/58.
 XX
 PT New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.
 XX
 PS Claim 1; SEQ ID NO 35521; 44bp; English.
 XX
 CC The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC segdata.uspto.gov/sequence.html?DocID=20030073623

XX	Sequence	422 BP; 95 A; 125 C; 121 G; 81 T; 0 U; 0 Other;
SQ	Alignment Scores:	
	Pred. No.:	1,82e-27
	Score:	324.00
	Percent Similarity:	78.95%
	Best Local Similarity:	54.39%
	Query Match:	33.68%
DB:		9
	US-10-627-571-2 (1-188) x ACH48309 (1-422)	0
QY	6 PheAaSerLysAaAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMetValSer	25
DB	77 TTCAGCTCAAGAGCCTGGCACTGCACAGCAGAGAAAGCTACTGAGTAAGATGGCGGCT	136
QY	26 LysSerIleAlaThrThrLeuIleAaPAspThrSerSerLysValLeuAaPGLuLeuTyr	45
DB	137 CGCTGTGGCTCATCTCTTCATAGATGAAACAGCAGTGAAGTGCTAGATGAGCTTAC	196
QY	46 ArgValThrArgGluTyrThrGlnAaAsnLysGlnAlaGluLysLysIleLysAaAsnLeu	65
DB	197 CGTGTCCTCAAGAGATACAGCAGCAGCCGAGCCAGGCCAGCGCGATCAAGACCTG	256
QY	66 IleLysThrValIleLysLeuAlaIleLysTyrArgAaAsnAaGlnPheAaGlnAaPGLu	85
DB	257 ATCAAAAGTGCATCAAGGTGGCTGTGCTGCACCGCAATGCTCTTTGGCCCAAGTAG	316
QY	86 LeuAlaLeuMetGlnLysPheLysLysLysValHisGlnLeuAlaMetThrValValSer	105
DB	317 CTGGCCCTGGCTAACCCGCTTTCGCGAAGAGTCGCGAGGCTGCATGACGGCATTAGC	376
QY	106 PheHisGlnValAspTyrThrPheAaPAspAaGlnValLeuSer	119
DB	377 TTTGGTGAAGTAGACTTCACTTCGAGGCTGCTGTCTGGCT	418
RESULT 35		
ACH76016		
ID	ACH76016 standard; DNA; 500 BP.	
XX	ACH76016;	
AC	ACH76016;	
DT	29-JUL-2004 (first entry)	
XX	Human genome derived single exon probe #9211.	
DE	Human; probe; ss; gene expression; single exon probe; microarray;	
XX	Human; probe; ss; gene expression; single exon probe; microarray;	
KW	alternative splicing event; genomic alteration.	
XX	Human; probe; ss; gene expression; single exon probe; microarray;	
OS	Homo sapiens.	
XX	US2003194704-A1.	
PN	16-OCT-2003.	
PD	03-APR-2002; 2002US-00029386.	
PF	03-APR-2002; 2002US-00029386.	
PR	03-APR-2002; 2002US-00029386.	
XX	(PENNY) PENN S G.	
PA	(RANK) RANK D R.	
XX	(HANZ) HANZEL D K.	
PI	Penn SG, Rank DR, Hanzel DK;	
XX	WPI, 2004-119264/12.	
DR	New human genome-derived single exon nucleic acid probes useful for human	
PT	gene expression analysis, for identifying or characterizing alternative	
PT	splicing events, for assessing genomic alterations or as tools for	
XX	surveying tissues.	

PS Claim 15, SEQ ID NO 9211, 80bp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
XX expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX

XX Sequence 500 BP; 164 A; 71 C; 107 G; 158 T; 0 U; 0 Other;

XX

XX Alignment Scores:

XX Pred. No: 4.25e-25 Length: 500

XX Score: 304.00 Matches: 55

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 31.60% Gaps: 0

XX Indels: 0

XX Gaps: 0

US-10-627-571-2 (1-188) x ACH76016 (1-500)

QY 134 GlnArgHisIeuThrAlaIysSerHisGlyArgValaAsnValPheAspHisPheSer 151

Db 3 CACGGCCACCTCACTGCCAAGCAACATGGAACGGGTAAATATGTTGGAACATTTTCA 62

QY 154 AspCysGluPheLeuAlaIleuLeuTyrAsnProPheGlyAsnPheLysProHisIleuGln 173

Db 63 GATTGGCAATTTTGGCTCCTCTGTATATCTTTTGGGAATTTTAAACCCCACTTCA 122

QY 174 LysLeuCysAspGlyIleAsnLysMetLeuAspGluIleuAsnIle 188

Db 123 AAACATATGATGTATCAACAAATGTTGGATGAAAGAACATTA 167

RESULT 36

AAH04721

ID AAH04721 standard; cDNA; 714 BP.

XX AAH04721;

XX

XX 26-JUN-2001 (first entry)

XX

XX Human cDNA clone (5'-primer) SEQ ID NO:1556.

XX

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

```

XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX PI WPI; 2001-318749/34.
XX DR
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 1; SEQ ID NO 1556; 2537bp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification; where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX SQ Sequence 714 BP; 113 A; 246 C; 203 G; 148 T; 0 U; 4 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,566-23 Length: 714
XX Score: 297.00 Matches: 50
XX Percent Similarity: 81.188 Conservative: 19
XX Best Local Similarity: 58.828 Mismatches: 16
XX Query Match: 30.354 Indels: 0
XX DB: 4 Gaps: 0
XX
XX US-10-627-571-2 (1-188) x AAH04721 (1-714)
XX
XX QY 104 ValSerPheHisGlnValAspTyrThrPheAspArgAsnValIeuSerArgIleuLeuAan 123
XX DB 6 GTCAGCTTCACACGAGTGAACCTTCGACCGCGCGTGGCCGCCGCCGAGCTGCCTC 65
XX
XX QY 124 GluGValArgGluMetLeuHisGlnIleIleGlnArgHisIleuThrAlaIysSerHisGly 143
XX DB 66 GAGTCCCGGACCTGCTGACCCAGGCGCGTGGTCCCGACCTGACCGCCAAAGTCCACGGCG 125
XX
XX QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTyrAan 163

```

```

DB 126 CGCATCAACACAGTTCCTGGCCACCTACGCACTGCTGCGCTACAGGC 185
QY 164 PropheGlyAsnPhelyProHisIeuGlnIlyLeuCyAspGlyTleAsnIyMetIeu 183
DB 186 CCCGCCGAGCCCTACCGCTCCACCTGCGCAGATCTCGAGGCGCTGCGCGATGCTG 245
QY 184 AspGluGluAanIle 188
DB 246 GACGAGGCGAGCCTC 260
RESULT 37
AAH18715
ID AAH18715 standard; cDNA; 1602 BP.
AC AAH18715;
XX
XX DT 26-JUN-2001 (first entry)
XX
XX DE Human cDNA sequence SEQ ID NO:18986.
XX
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX OS Homo sapiens.
XX
XX XX EP1074617-A2.
XX
XX PD 07-FEB-2001.
XX
XX PF 28-JUL-2000; 2000EP-00116126.
XX
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX
XX PA (HELI-) HELIX RES INST.
XX
XX PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX PI WPI; 2001-318749/34.
XX
XX DR
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 8; SEQ ID NO 18986; 2537bp + Sequence Listing; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification; where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention

```


PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251899P.
 PR 08-DEC-2000; 2000US-0251907P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 DR P-PSDB; AAM83841.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Claim 1; SEQ ID NO 1682; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169

CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 441 BP; 99 A; 119 C; 127 G; 89 T; 0 U; 7 Other;
 Alignment Scores:
 Pred. No.: 1,47e-22 Length: 441
 Score: 281.00 Matches: 60
 Percent Similarity: 74.36% Conservative: 27
 Best Local Similarity: 51.28% Mismatches: 24
 Query Match: 29.21% Indels: 7
 DB: 4 Gaps: 1
 US-10-627-571-2 (1-188) x AAK56622 (1-441)
 QY 6 PheAsnSerLyAsnLeuAlaValGlnAlaGlnLyLeuGlyLeuGlyLeuValSer 25
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 89 TTCAGCTCAAGAGCGCTGCGACCTGCGACGAGGAGAAAGCTACTAGTAGTGCGGGT 148
 QY 26 LySeriLeaLThrThriLeuLeaAspTrSerSerGluValLeuAspGluLeuTyr 45
 ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 149 CGCTGTGGCTCAGCTCTCATAGATGAGACAGACAGTGAAGTGTGCTAGATGAGCTTAC 208
 QY 46 ArgValThrArgGluTyrThrGlnAsnLyLeuGlnAlaGlnLyLeuValLeuValSer 65
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 209 CGTGTCCAGAGAGTACCGCAGCGCGCCGCGCCAGCGCGGTGATCAGAGACCTG 268
 QY 66 IleLyThrValIleLySerLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnAspGlu 85
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 269 ATCANAGTGGCATCAAGGTGCTGTCTGCACCGCATGTCTCTTGCGCCAGTGA 328
 QY 86 LeuAlaLeuMetGlnLysPheLysLysValLHsGlnLeuAlaMetThrValValSer 105
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 329 CTGGCCCTGCTTACCGCGCTTCCGCAAGARCTGCGCAAGGTGCGCATGAGCATTTAGAC 388
 QY 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeu 122
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 389 TTGGT-GAAGTGAATTTCACT-----TCGAAGCTGCTG 420
 RESULT 39
 AAC10352
 ID AAC10352 standard; cDNA; 340 BP.
 AC AAC10352;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 14427.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 PN EP1033401-A2.
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 XX
 PA (GENST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 1; SEQ ID NO 14427; 71pp + Sequence Listing; English.
 XX

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNA or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors

XX SQ Sequence 340 BP; 92 A; 79 C; 85 G; 79 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	3,44e-19	Length:	340
Score:	250.00	Matches:	52
Percent Similarity:	96.36%	Conservative:	1
Best Local Similarity:	94.55%	Mismatches:	2
Query Match:	25.99%	Indels:	0
DB:	3	Gaps:	0

US-10-627-571-2 (1-188) x AAC10352 (1-340)

QY 2 AlAThAspYAlPheAsnSerLYsAsnLeuAlaValGlnAgluLYsLYleuGly 21
DB 174 GCACAGATGCTCTTAATTCAGAAAACCTGGCGCKTCAGGACGCAAAAGATTTGGT 233

QY 22 LyMeMetValSerLYsSerLYleAlaThrThreuLYleAspAspThrSerSerGluValLeu 41
DB 234 AAATGTGTGTCGAATCATGCGCACCACTTAATAGACGACCAAGTAGTGAGTCTG 293

QY 42 AspGluLeuTYrArgValThrArgGluTYrThrGlnAsnLYsLYs 56
DB 294 GATGAGCTCTACAGAGTGACGAGGAGTACCCCAAAACAGAG 338

RESULT 40
AA192435/c
ID AA192435 standard; cDNA; 400 BP.

XX AC AA192435;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 12495.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PA 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR P-PSDB; AAO12504.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 1; SEQ ID NO 12495; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AA19941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 400 BP; 80 A; 122 C; 114 G; 84 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.2e-14	Length:	400
Score:	206.00	Matches:	38
Percent Similarity:	85.25%	Conservative:	14
Best Local Similarity:	62.30%	Mismatches:	9
Query Match:	21.41%	Indels:	0
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x AA192435 (1-400)

QY 2 AlAThAspYAlPheAsnSerLYsAsnLeuAlaValGlnAgluLYsLYleuGly 21
DB 184 GCCATGGACACCTTTCAGACCAAGAGCTGGCTCTCAGGCGGAGAAAGCTCTGAGT 125

QY 22 LyMeMetValSerLYsSerLYleAlaThrThreuLYleAspAspThrSerSerGluValLeu 41
DB 124 AAGATGGCTTCAAGGACGTGGCGGCTGTGTGTGATGACCCAGCATGTGAGTCTG 65

QY 42 AspGluLeuTYrArgValThrArgGluTYrThrGlnAsnLYsLYs 61
DB 64 GATGAGCTGTACCGGCCACGAGGAGTTCACGCGGACGCCGAGAGGCCAGAGG 5

QY 62 Ile 62
DB 4 CTC 2

RESULT 41
AAC10359
ID AAC10359 standard; cDNA; 252 BP.

XX AC AAC10359;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 14434.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX PN BP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-00200610.

XX PR 26-FEB-1999; 99US-0122487P.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 14434; 71bp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A⁺ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 252 BP; 57 A; 88 C; 60 G; 46 T; 0 U; 1 Other;
XX
Alignment Scores:
Pred. No.: 7 92e-10 Length: 252
Score: 166.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.26% Indels: 0
Gaps: 0
DB: 3
US-10-627-571-2 (1-188) x AAC10359 (1-252)
QY 1 MetAlaThrAspValPheAsnSerLySAstLeuAlaValGlnAlaGlnLysIleLeu 20
DB 148 ATGGCCACAGATGCTTTTAATTCGAAAACCTGCGCGTTCAGGCACAAAAGAGACTTC 207
QY 21 GlyLysMetValSerLysSerIleAlaThrThreulIleAspAsp 35
DB 208 GGTAAATGGTGTCCAAATCCATCGCCACCACTTAATAGACGAC 252
RESULT 42
ACH78513
ID ACH78513 standard; DNA; 543 BP.
XX
XX ACH78513;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #11708.
XX
XX Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
XX US2003194704-A1.
XX
XX 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX

XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 15; SEQ ID NO 11708; 80bp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 543 BP; 137 A; 128 C; 119 G; 159 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.4e-07 Length: 543
Score: 148.00 Matches: 24
Percent Similarity: 88.64% Conservative: 15
Best Local Similarity: 54.55% Mismatches: 5
Query Match: 15.38% Indels: 0
Gaps: 0
DB: 12
US-10-627-571-2 (1-188) x ACH78513 (1-543)
QY 143 GlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyr 162
DB 3 GGCGCATCAACACGCTTTTAACCACTTGGCGAGTGGAGTCCCTCCACCCCTCAT 62
QY 163 AsnProPheGlyAsnPhelySProHisLeuGlnLysLeuCyAspGlyIleAsnLysMet 182
DB 63 AGCTCGATGAGACAGCTGAGCGCCCAACTCAAGAGATTGTGAAGAGATCATTAAGTTG 122
QY 183 LeuAspGluGlu 186
DB 123 CTAGATGAGAAA 134
RESULT 43
ADP04796
ID ADP04796 standard; cDNA; 1329 BP.
XX
XX ADP04796;
XX

XX 29-JUL-2004 (first entry)
DT Sea squirt CDNA with tissue specific expression in development Seq 391.
DE gene; ss; sea squirt; regeneration medicine; gene therapy;
XX cell proliferation; differentiation; reproduction;
KM environmental measurement; water survey.
XX Clona intestinalis.
OS JP2004057129-A.
PN 26-FEB-2004.
PD 31-JUL-2002; 2002JP-00222593.
PF 31-JUL-2002; 2002JP-00222593.
PR 31-JUL-2002; 2002JP-00222593.
XX (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
PA WPI; 2004-287079/27.
DR P-PSDB; ADP04797.
PT Novel gene cluster which is specifically expressed in tissue or organ
PT during developmental phase of sea squirt, useful for elucidation of
PT mechanism of development of tissue or organ of sea squirt.
XX Claim 2; SEQ ID NO 391; 1846bp; Japanese.
XX
CC This invention relates to novel genes and the encoded proteins thereof
CC that are derived from the sea squirt Clona intestinalis. Specifically, it
CC refers to those genes that are expressed in the tissues or organs of the
CC sea squirt during its developmental phase. The present invention
CC describes the identification of these genes as useful for elucidation of
CC the mechanism of development and hence for developing regeneration of
CC medicines and gene therapy techniques. Accordingly, they can be used in
CC the research of various genetic diseases, as well as the analysis of cell
CC proliferation, differentiation and reproduction. Furthermore, such
CC compositions can be useful for environmental measurements and water
CC surveys, particularly for sea water surveys, and also for the preparation
CC of transformed sea squirt for improving edibility of sea squirt such as
CC Halocynthia roretzi. This polynucleotide sequence is a sea squirt CDNA
CC sequence that exhibits tissue specific expression during development.
CC given in an exemplification of the invention.
XX
SQ Sequence 1329 BP; 596 A; 183 C; 280 G; 270 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0.0456 Length: 1329
Score: 106.00 Matches: 45
Percent Similarity: 44.51% Conservative: 36
Best Local Similarity: 24.73% Mismatches: 59
Query Match: 11.02% Indels: 42
DB: 12 Gaps: 8
US-10-627-571-2 (1-188) x ADP04796 (1-1329)
QY AAsenSetlysaSnleuAlaValGlnAlaGlnlyLeuGlylyMetValSerlyS 26
DB 242 AATGAAGAAAGAACTCGCTGCAAGAGCCGAGAAACAA----- 277
QY 27 SerilAlaIaThrIleuIleAaPThrSerSerGluValIleuAaPgluLeuTyraS 46
DB 278 -----AAGAAATTAATAAGATGAAGAAAGAAAGTTGAGAGGAG-----AGG 319
QY 47 ValThraGluIuTyThrGlnAaenlySlyGluAlaGlnlySlylelysaSnleuIle 66
DB 320 GTr---CGTAAACAACTGAAGCAAGAAAGAAAGAAAGAAAGAAAGAAATGCTTAA 376
QY 67 LyEThrValIlelySleuAlaIleLeuTyraGAsnaSnInPhe-----Aa 82
DB 377 AAGAAAGCAAAATTTAAAGAGCTATGTGAAGATCATATTTCTTCTGATTTCTGAA 436

QY 83 GlnaSPGluLeuAlaIeumetGluIySphelySlySlyValHIsGlnleuAlaMetThr 102
DB 437 GAAGAAAGGTTAAATGTGATGAA-----CAATGTATCACTTGTTGTGAA 484
QY 103 Val-----ValSerPheHIsGlnValaPlyTyThrPheaSP----- 114
DB 485 GTCAATCTTATTTTCATTGCAAGATTAACAAATGCAATAGTCAGCAATTTACTGAA 544
QY 115 -----ArgaSnValIeuserArgleuLeuAaenGluCyArgGluMetLeuHIsGlnIle 132
DB 545 AGTGCAGAGATATATATTTGACAAAGCTTCACAACTTCGCAACAAGTTGAACAAG 604
QY 133 lIeGlnaRgHIsleuThraIaIySserHIsGlyArgValaSnAaSnValPheaSP----- 150
DB 605 AAACAGAAACATTTGGACAGCTTCACAGTCAAAAGATTAACCTGTGATGAGATTTCA 664
QY 151 -----HisPheSerAaPcyGluPheleuAlaIaIeutyraS 163
DB 665 AGTGCAGAGAGAAACATGAGTTACAGTATTAATCTTGATTAAGACATCAAT 724
QY 164 ProPhe 165
DB 725 CTCCTT 730
RESULT 44
ID AAA70229 standard; DNA; 2943 BP.
XX AAA70229;
AC
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:362.
XX
KM Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX antimalarial; malaria; protozoasidae; infection; insecticide; ds.
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US026796.
PR 05-NOV-1998; 98US-0107131P.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
PT Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P. falciparum infection.
XX
PS Disclosure; Page 549-550; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (i) nucleotide sequences (ii) encoding (i); and (2)
CC vaccines against P. falciparum infection comprising (i) or (ii). (1) and
CC (ii) are useful for the development of vaccines against P. falciparum
CC infection. (i) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (i), are useful in the detection
CC of infection with P. falciparum. Furthermore, (i) (especially when they
CC are refined or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent P. falciparum infection, or they can be used


```
Qy 61 LysIleLysAsnLeuIleuThrValIleuLysLeuAlaIleuLeuTyrArgAsnAngln 80
Db 224 AGAGAGAGAGAACTG-----GAGTTGAATGTGTATCAGAAAGCAACCTG 268
Qy 81 PheAsnGlnAspGluLeuAlaLeuMetGlnLys-----PheLysLysLysValHisGln 98
Db 269 AATGTGCATTTAAATTCTTCTCTCTCGAGAGAGATGATCTTAAAGAGCTTCATGCT 328
Qy 99 LeuAlaMetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeu 118
Db 329 ATTCAAAAGGAATACAGTGAATTCAGTCC-----AATCTGATGAGAAAGTGCT 379
Qy 119 Ser-----ArgLeuLeuAsnGlnCysArgGluMetLeuHisGlnIleIleGlnArgHis 136
Db 380 TCAAGATCGAAGCTGTGGGGGACCAAGAGAGAGACTCACCCAGCTTGAAGAACAACTT 439
Qy 137 LeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGlu 156
Db 440 GGCACTGCCTTAAGTGAAGCAAGTAAATGAAGTG----- 475
Qy 157 PheLeuAlaAlaLeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCys 176
Db 476 CTAATGCTGATCTGACTCGAGAAAAGAAAACCTTAGAGAAATGATGATGCTGAGCTG 535
Qy 177 AspGlyTLeuAsnLysMetLeuAspGlu 185
Db 536 GACAAATGTAAACAAGTTAAAGCAAGAG 562
```

Search completed: July 28, 2005, 19:22:51
Job time : 563 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 19:01:22 ; Search time 622 Seconds

(Without alignments)
1956.275 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962

Sequence: 1 MARDVRSKRLAVQAQKIL.....KPHLQKLCDCGINKMLDEBNT 188

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Egapop 6.0 , Egapext 7.0

Delop 6.0 , Delext 7.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xljh

-O=/cg2_1/USPTO.spool/h/US10627571/runat.27072005.154722.27521/app.query.fasta_1.327

-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=tblsum62

-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR_SCORE=spct -THR_MAX=100

-THR_MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0

-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELTOP=6 -DELEXT=7

Database :

Published Applications_NA:*

1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cg2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*

3: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

5: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

6: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

7: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

8: /cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*

9: /cg2_6/ptodata/2/pubpna/US12_NEW_PUB.seq:*

10: /cg2_6/ptodata/2/pubpna/US13_NEW_PUB.seq:*

11: /cg2_6/ptodata/2/pubpna/US14_NEW_PUB.seq:*

12: /cg2_6/ptodata/2/pubpna/US15_NEW_PUB.seq:*

13: /cg2_6/ptodata/2/pubpna/US16_NEW_PUB.seq:*

14: /cg2_6/ptodata/2/pubpna/US17_NEW_PUB.seq:*

15: /cg2_6/ptodata/2/pubpna/US18_NEW_PUB.seq:*

16: /cg2_6/ptodata/2/pubpna/US19_NEW_PUB.seq:*

17: /cg2_6/ptodata/2/pubpna/US20_NEW_PUB.seq:*

18: /cg2_6/ptodata/2/pubpna/US21_NEW_PUB.seq:*

19: /cg2_6/ptodata/2/pubpna/US22_NEW_PUB.seq:*

20: /cg2_6/ptodata/2/pubpna/US23_NEW_PUB.seq:*

21: /cg2_6/ptodata/2/pubpna/US24_NEW_PUB.seq:*

22: /cg2_6/ptodata/2/pubpna/US25_NEW_PUB.seq:*

23: /cg2_6/ptodata/2/pubpna/US26_NEW_PUB.seq:*

24: /cg2_6/ptodata/2/pubpna/US27_NEW_PUB.seq:*

25: /cg2_6/ptodata/2/pubpna/US28_NEW_PUB.seq:*

26: /cg2_6/ptodata/2/pubpna/US29_NEW_PUB.seq:*

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	954	99.2	1892	22 US-10-491-545A-48	Sequence 48, Appl
2	954	99.2	1915	18 US-10-627-571-1	Sequence 1, Appl1
3	950	98.8	1814	19 US-10-755-889-215	Sequence 215, Appl
4	950	98.8	2003	19 US-10-755-889-217	Sequence 217, Appl
5	950	98.8	2034	13 US-10-087-192-497	Sequence 497, Appl
6	950	98.8	2081	19 US-10-755-889-39	Sequence 39, Appl
7	950	98.8	58723	13 US-10-087-192-496	Sequence 496, Appl
8	948	98.5	1943	14 US-10-097-065-98	Sequence 98, Appl
9	920	95.6	2087	13 US-10-087-192-494	Sequence 94, Appl
10	920	95.6	62231	13 US-10-087-192-493	Sequence 493, Appl
11	920	95.6	544	16 US-10-029-386-22314	Sequence 22314, A
12	914	95.0	544	16 US-10-029-386-22314	Sequence 22314, A
13	914	95.0	544	16 US-10-029-386-22314	Sequence 22314, A
14	610	63.4	527	16 US-10-094-466-43	Sequence 43, Appl
15	593	61.6	619	17 US-10-094-466-41	Sequence 41, Appl
16	593	61.6	2108	18 US-10-416-314-70	Sequence 70, Appl
17	591	61.4	1188	9 US-09-816-828-5	Sequence 5, Appl1
18	566	58.8	1389	21 US-10-764-420-2116	Sequence 2116, Ap
19	527	54.8	1113	9 US-09-822-849A-81	Sequence 81, Appl
20	527	54.8	1165	21 US-10-363-374-15	Sequence 15, Appl
21	527	54.8	1175	18 US-10-302-172-537	Sequence 537, Appl
22	527	54.8	1268	21 US-10-959-539-59	Sequence 59, Appl
23	504	52.4	287	9 US-09-796-692-3207	Sequence 3207, Ap
24	504	52.4	287	14 US-10-040-862-3207	Sequence 3207, Ap
25	504	52.4	287	17 US-10-057-475B-3207	Sequence 3207, Ap
26	504	52.4	287	17 US-10-154-884B-3207	Sequence 3207, Ap
27	504	52.4	287	19 US-10-764-324-3207	Sequence 3207, Ap
28	499	51.9	287	9 US-09-796-692-2801	Sequence 2801, Ap
29	499	51.9	287	14 US-10-040-862-2801	Sequence 2801, Ap
30	499	51.9	287	17 US-10-057-475B-2801	Sequence 2801, Ap
31	499	51.9	287	17 US-10-154-884B-2801	Sequence 2801, Ap
32	499	51.9	287	19 US-10-764-324-2801	Sequence 2801, Ap
33	484	50.3	288	9 US-09-796-692-5724	Sequence 5724, Ap
34	484	50.3	288	14 US-10-040-862-5724	Sequence 5724, Ap
35	484	50.3	288	17 US-10-057-475B-5724	Sequence 5724, Ap
36	484	50.3	288	17 US-10-154-884B-5724	Sequence 5724, Ap
37	484	50.3	422	10 US-09-918-999-35521	Sequence 35521, A
38	324	33.7	422	10 US-09-918-999-35521	Sequence 35521, A
39	304	31.6	500	16 US-10-029-386-9211	Sequence 9211, A
40	189	19.6	565	22 US-10-972-079-232	Sequence 232, Appl
41	155	16.1	433	17 US-10-191-803-330	Sequence 330, Appl
42	148	15.4	533	16 US-10-029-386-11708	Sequence 11708, A
43	109	11.3	538	9 US-09-796-692-4917	Sequence 4917, Ap
44	109	11.3	538	14 US-10-040-862-4917	Sequence 4917, Ap
45	109	11.3	538	17 US-10-057-475B-4917	Sequence 4917, Ap
46	109	11.3	538	17 US-10-154-884B-4917	Sequence 4917, Ap
47	109	11.3	538	19 US-10-764-324-4917	Sequence 4917, Ap
48	95	9.9	638	9 US-09-974-300-6328	Sequence 6328, Ap
49	93.5	9.7	2343	17 US-10-282-122A-10827	Sequence 10827, A
50	92.5	9.6	2501	14 US-10-198-884B-13958	Sequence 13958, A
51	91	9.5	1305	21 US-10-031-070-2799	Sequence 279, Appl
52	90.5	9.4	1359	17 US-10-282-122A-35257	Sequence 35257, A
53	90.5	9.4	4232	9 US-09-745-763-167	Sequence 167, Appl
54	90.5	9.4	4232	10 US-09-814-355-21754	Sequence 21754, A
55	90.5	9.4	4428	17 US-10-295-021-445	Sequence 445, Appl
56	90	9.4	1887	16 US-10-349-680-133	Sequence 133, Appl
57	89.5	9.3	1142	13 US-10-082-018-8	Sequence 8, Appl1
58	89.5	9.3	1142	21 US-10-949-978-8	Sequence 8, Appl1
59	88.5	9.2	1700	9 US-09-954-531-623	Sequence 623, Appl
60	88.5	9.2	1700	17 US-10-305-720-1081	Sequence 1081, Ap
61	88.5	9.2	1700	18 US-10-641-643-898	Sequence 898, Appl
62	88.5	9.2	1700	21 US-10-843-641A-1690	Sequence 1690, Appl
63	88	9.1	1109	19 US-10-767-701-9360	Sequence 9360, Ap
64	88	9.1	1638	9 US-09-976-740-46	Sequence 46, Appl
65	88	9.1	1638	13 US-10-023-529-46	Sequence 46, Appl

Pred. No. is the number of results predicted by chance to have a

```

66      88      9.1 1638 13 US-10-023-523-46 Sequence 46, Appl
67      88      9.1 1638 17 US-10-616-187-44 Sequence 46, Appl
68      88      9.1 1638 17 US-10-671-242-46 Sequence 46, Appl
69      88      9.1 1793 20 US-10-473-974-165 Sequence 165, Appl
70      88      9.1 2523 18 US-10-276-774-784 Sequence 784, Appl
71      88      9.1 4697 9 US-09-962-055-17 Sequence 17, Appl
72      88      9.1 4697 9 US-09-976-740-17 Sequence 17, Appl
73      88      9.1 4697 13 US-10-023-529-17 Sequence 17, Appl
74      88      9.1 4697 13 US-10-023-523-17 Sequence 17, Appl
75      88      9.1 4697 17 US-10-616-187-17 Sequence 17, Appl
76      88      9.1 4697 18 US-10-671-242-17 Sequence 17, Appl
77      87.5     9.1 1503 9 US-09-938-842A-1473 Sequence 1473, Ap
78      87.5     9.1 2341 9 US-09-938-842A-1473 Sequence 1473, Ap
79      87.5     9.1 3349 10 US-09-881-752A-147 Sequence 147, App
80      87.5     9.1 3349 10 US-09-952-267-2 Sequence 2, Appl
81      87.5     9.1 3349 22 US-10-872-768-2 Sequence 2, Appl
82      87.5     9.1 3349 22 US-10-872-768-2 Sequence 2, Appl
83      87.5     9.1 4155 17 US-10-062-674-2109 Sequence 2109, Ap
84      87.5     9.1 4722 9 US-09-962-055-14 Sequence 14, Appl
85      87.5     9.1 4722 9 US-09-976-740-14 Sequence 14, Appl
86      87.5     9.1 4722 13 US-10-023-529-14 Sequence 14, Appl
87      87.5     9.1 4722 13 US-10-023-523-14 Sequence 14, Appl
88      87.5     9.1 4722 17 US-10-616-187-14 Sequence 14, Appl
89      87.5     9.1 4722 18 US-10-671-242-14 Sequence 14, Appl
90      87.5     9.0 1042 17 US-10-426-776-47 Sequence 47, Appl
91      87       9.0 3249 19 US-10-437-963-64679 Sequence 64679, A
92      87       9.0 22507 16 US-10-085-959-53 Sequence 53, Appl
93      86.5     9.0 744 17 US-10-282-122A-18136 Sequence 18136, A
94      86.5     9.0 1604 18 US-10-425-114-20487 Sequence 20487, A
95      86.5     9.0 1667 20 US-10-425-115-178606 Sequence 178606,
96      86.5     9.0 2052 9 US-09-792-630-40 Sequence 40, Appl
97      86.5     9.0 2052 10 US-09-953-351-40 Sequence 40, Appl
98      86.5     9.0 2052 13 US-10-080-376-40 Sequence 40, Appl
99      86.5     9.0 2052 14 US-10-082-671-46 Sequence 46, Appl
100     86.5     9.0 2052 14 US-10-097-100-40 Sequence 40, Appl

```

ALIGNMENTS

```

RESULT 1
US-10-491-545A-48
; Sequence 48, Application US/10491545A
; Publication No. US20050130117A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Peter
; APPLICANT: Li, Congfen
; APPLICANT: Liao, X. Charlene
; APPLICANT: Masuda, Esteban
; APPLICANT: Pardo, Jorge
; APPLICANT: Zhao, Haoran
; APPLICANT: Rigol Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modulators of Lymphocyte Activation and Migration
; FILE REFERENCE: 021044-000330US
; CURRENT APPLICATION NUMBER: US/10/491, 545A
; CURRENT FILING DATE: 2004-04-01
; PRIOR APPLICATION NUMBER: US 60/327,212
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: WO PCT/US02/31618
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human TNF-induced protein GG2-1 mRNA,
US-10-491-545A-48

```

Alignment Scores:

```

Score: 5.74e-97 Length: 1892
Pred. No.: 954.00 Matches: 187

```

```

Percent Similarity: 99.47%
Best Local Similarity: 99.47%
Query Match: 99.17%
DB: 22 Gaps: 0
US-10-627-571-2 (1-188) x US-10-491-545A-48 (1-1892)
1 MetAlaThAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
98 ATGGCCACGATGTCCTTATTCCTTATTCCTTATTCCTTATTCCTTATTCCTTATTC 157
21 GlyIysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerGluVal 40
158 GGTAAATGTCCTTCAAAATCCATCCGACCACTTATTCAGACGACCAAGTGTGGG 217
41 LeuAspGluLeuThrArgValThrArgGluThrGlnAsnIleLysGluAlaGluLys 60
218 CTGGATGAGCTCTACAGACTGACGAGGAGTACACCCAAACAAAGAGGACAGGAAG 277
61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLysThrArgAsnGln 80
278 ATCATCAAGACCTCATCAAGACCTCATCAAGACCTCATCAAGACCTCATCAAGAC 337
81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAla 100
338 TTTAATCAAGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 397
101 MetThrValValSerPheIleGlnValAspThrPheAspArgAsnValLeuSerArg 120
398 ATGACCGTGGTCAATTCATCAAGATGATGATGATGATGATGATGATGATGATGAT 457
121 LeuLeuAsnGluCysArgGluMetLeuIleGlnIleIleGlnArgHisLeuThrAlaLys 140
458 CTGTTAAAGATGACGAGAGATGCTGACCAAAATCATTACGCCCACTCTGCCAAG 517
141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
518 TCACATGACGAGGATTAATATGATCTTTGATCATTTTTCAGATTGGAATTTTGGCTGC 577
161 LeuThrAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGluLeu 180
578 TTGTAAATCCTTTTGGGAATTTTAAACCCCTTAAACAAACATATGATGATGATCAAC 637
181 LysMetLeuAspGluGluAsnIle 188
638 AAAATCTTGATCAAGACCACTA 661

```

RESULT 2

```

US-10-627-571-1
; Sequence 1, Application US/10627571
; Publication No. US20040082771A1
; GENERAL INFORMATION:
; APPLICANT: KASID, Usha N.
; APPLICANT: KUMAR, Deepak
; APPLICANT: GOKHALE, Prafulla
; APPLICANT: AHMAD, Imran
; TITLE OF INVENTION: ANTI-APOTOPIC GENE SCC-S2 AND DIAGNOSTIC AND THERAPEUTIC USBS
; FILE REFERENCE: 223316
; CURRENT APPLICATION NUMBER: US/10/627,571
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: US 60/264,062
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US02/02212
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```



```
/ ORGANISM: Homo sapiens
US-10-755-889-217

Alignment Scores:
Pred. No.: 1,77e-96      Length: 2003
Score: 950.00           Matches: 186
Percent Similarity: 99.47%  Conservative: 1
Best Local Similarity: 98.94%  Mismatches: 1
Query Match: 98.75%      Indels: 0
DB: 19                  Gaps: 0

US-10-627-571-2 (1-188) x US-10-755-889-217 (1-2003)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeu 20
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 204 GTGGCCACAGAGTCTTTAATTCCAAAACCTGGCCGTTCAGGCACAAAGAGATCTTG 263

QY 21 GlyLysMetValSerLysSerLleAlaThrThrLeuLleAspAspThrSerSerGluVal 40
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 264 GGTAAATGCTGCCAATTCATCGCCACCACTTATAGACGACACAAAGTAGAGTG 323

QY 41 LeuAspGluLeuTyrrArgValThrArgLutTyrrThrGlnAsnLysLysGluAlaGluLys 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 324 CTGGATGACCTTCAAGAGTACAGGAGGTACACCAAAACAAAGAGGAGGAGGAAG 383

QY 61 LysLleLysAsnLeuLleLysThrValLleLysLeuAlaLleLeuTyrrArgAsnAsnGln 80
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 384 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCG 443

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 444 TTTAATCAAGATGAGCTGATGATGAGAAATTTAGAGAAATGTCATCAGTTGCT 503

QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 504 ATGACCGTGTCAGATTTCATCAAGTGAATTAACCTTTGACCGGATGTTATCCAGG 563

QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnLleLleGlnArgHisLeuThrAlaLys 140
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 564 CTGTTAAATGAATGACGAGAGAGATGCTGCACCAATCATTCAGGCCACTCTGCCAAG 623

QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 624 TCACATGACGCGGTTAATATGTTGTTGATCATTTTTCAGATGTGAATTTTGGCTGCC 683

QY 161 LeuTyrrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuAspGluLysLeuAla 180
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 684 TTGTAATATCCTTTGGGAATTTTAAACCCCACTTACAAAACTATGTATGTATCAAC 743

QY 181 LysMetLeuAspGluGluAsnLle 188
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 744 AAAATGTTGATGAAGAGAACATA 767

RESULT 5
US-10-087-192-497
/ Sequence 497, Application US/10087192
/ Publication No. US20020182586A1
/ GENERAL INFORMATION:
/ APPLICANT: Morris, David W.
/ APPLICANT: Engelhard, Eric K.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
/ FILE REFERENCE: 529452000122
/ CURRENT APPLICATION NUMBER: US/10/087,192
/ PRIOR FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ PRIOR FILING DATE: 2001-03-02
/ NUMBER OF SEQ ID NOS: 2059
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 497
/ LENGTH: 2034
```

```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-087-192-497

Alignment Scores:
Pred. No.: 1,81e-96      Length: 2034
Score: 950.00           Matches: 186
Percent Similarity: 99.47%  Conservative: 1
Best Local Similarity: 98.94%  Mismatches: 1
Query Match: 98.75%      Indels: 0
DB: 13                  Gaps: 0

US-10-627-571-2 (1-188) x US-10-087-192-497 (1-2034)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeu 20
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 248 GTGGCCACAGAGTCTTTAATTCCAAAACCTGGCCGTTCAGGCACAAAGAGATCTTG 307

QY 21 GlyLysMetValSerLysSerLleAlaThrThrLeuLleAspAspThrSerSerGluVal 40
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 308 GGTAAATGCTGCCAATTCATCGCCACCACTTATAGACGACACAAAGTAGAGTG 367

QY 41 LeuAspGluLeuTyrrArgValThrArgLutTyrrThrGlnAsnLysLysGluAlaGluLys 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 368 CTGGATGAGCTTCAAGAGTGAACGAGGTACACCAAAACAAAGAGGAGGAGGAAG 427

QY 61 LysLleLysAsnLeuLleLysThrValLleLysLeuAlaLleLeuTyrrArgAsnAsnGln 80
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 428 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCG 487

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 488 TTTAATCAAGATGAGCTGATGATGAGAAATTTAGAGAAATGTCATCAGTTGCT 547

QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 548 ATGACCGTGTCAGATTTCATCAGCTGATGATTAACCTTTGACCGGATGTTATCCAGG 607

QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnLleLleGlnArgHisLeuThrAlaLys 140
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 608 CTGTTAAATGAATGACGAGAGAGATGCTGCACCAATCATTCAGGCCACTCTGCCAAG 667

QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 668 TCACATGACGCGGTTAATATGTTGTTGATCATTTTTCAGATGTGAATTTTGGCTGCC 727

QY 161 LeuTyrrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGluLysLeuAla 180
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 728 TTGTAATATCCTTTGGGAATTTTAAACCCCACTTACAAAACTATGTATGTATCAAC 787

QY 181 LysMetLeuAspGluGluAsnLle 188
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 788 AAAATGTTGATGAAGAGAACATA 811

RESULT 6
US-10-755-889-39
/ Sequence 39, Application US/10755889
/ Publication No. US20040171823A1
/ GENERAL INFORMATION:
/ APPLICANT: Bristol-Myers Squibb Company
/ APPLICANT: Bristol-Myers Squibb Company
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
/ FILE REFERENCE: D0284 NP
/ CURRENT APPLICATION NUMBER: US/10/755,889
/ PRIOR FILING DATE: 2004-01-13
/ PRIOR APPLICATION NUMBER: U.S. 60/440,068
/ PRIOR FILING DATE: 2003-01-14
/ PRIOR APPLICATION NUMBER: U.S. 60/469,757
/ PRIOR FILING DATE: 2003-05-12
/ NUMBER OF SEQ ID NOS: 823
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 39
/ LENGTH: 2081
```

TYPE: DNA
ORGANISM: Homo sapiens
US-10-755-889-39

Alignment Scores:

Pred. No.:	1,876-96	Length:	2081
Score:	950.00	Matches:	186
Percent Similarity:	99.47%	Conservative:	1
Best Local Similarity:	98.94%	Mismatches:	1
Query Match:	98.75%	Indels:	0
DB:	19	Gaps:	0

US-10-627-571-2 (1-188) x US-10-755-889-39 (1-2081)

```

QY      1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnIaGlnIySlySllEleu 20
      ::::
DB      267 GTGGCCACAGATGCTTTTAATCCAAAACCTGGCGCTCAGGCACAAAAGAGATCTTG 326

QY      21 GlyLyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      |||||
DB      327 GGTAAATGAGTGTCCAAATCCATCGCCACCACTTAATAGACACCAAGTACTGAGGTG 386

QY      41 LeuAspGluLeuTyraArgValThrArgGluTyThrGlnAsnLySlySgluaIaGluLyS 60
      |||||
DB      387 CTGGATGAGCTCTACAGAGTGAACGAGGAGTACACCCAAAACAAAGAGGAGGAGAG 446

QY      61 LysIleLyAsnLeuIleLySthrValIleLySLeuAlaIleLeuTyraArgAsnGln 80
      |||||
DB      447 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAG 506

QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLySPhelySlyValHisGlnLeuAla 100
      |||||
DB      507 TTTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGTTCATCAGCTTGCT 566

QY      101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
      |||||
DB      567 ATGACCGTGTGCTGATTCATCAGTGGTGAATTATACCTTTGACCGAATGTGTTATCCAG 626

QY      121 LeuLeuAsnGluCyAsnArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||||
DB      627 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCCACCTCACTGCCAAG 686

QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
      |||||
DB      687 TCACATGAGACGGGTATATATGTTGATGATTTTTCAGATTGTGAATTTTGGCTGCC 746

QY      161 LeuTyraAsnProPheGlyAsnPhelySProHisLeuGlnLySLeuCyAspGlyIleAsn 180
      |||||
DB      747 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGTATCAAC 806

QY      181 LysMetLeuAspGluGluAsnIle 188
      |||||
DB      807 AAAATGTTGATGAAGAAACATA 830

RESULT 7
US-10-087-192-496
; Sequence 496, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496

```

LENGTH: 58723
TYPE: DNA
ORGANISM: Homo sapiens
US-10-087-192-496

Alignment Scores:

Pred. No.:	2,866-94	Length:	58723
Score:	950.00	Matches:	186
Percent Similarity:	99.47%	Conservative:	1
Best Local Similarity:	98.94%	Mismatches:	1
Query Match:	98.75%	Indels:	0
DB:	13	Gaps:	0

US-10-627-571-2 (1-188) x US-10-087-192-496 (1-58723)

```

QY      1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnIaGlnIySlySllEleu 20
      ::::
DB      46937 GTGGCCACAGATGCTTTTAATCCAAAACCTGGCGCTCAGGCACAAAAGAGATCTTG 46996

QY      21 GlyLyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      |||||
DB      46997 GGTAAATGAGTGTCCAAATCCATCGCCACCACTTAATAGACACCAAGTACTGAGGTG 47056

QY      41 LeuAspGluLeuTyraArgValThrArgGluTyThrGlnAsnLySlySgluaIaGluLyS 60
      |||||
DB      47057 CTGGATGAGCTCTACAGAGTGAACGAGGAGTACACCCAAAACAAAGAGGAGGAGAG 47116

QY      61 LysIleLyAsnLeuIleLySthrValIleLySLeuAlaIleLeuTyraArgAsnGln 80
      |||||
DB      47117 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAG 47176

QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLySPhelySlyValHisGlnLeuAla 100
      |||||
DB      47177 TTTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGTTCATCAGCTTGCT 47236

QY      101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
      |||||
DB      47237 ATGACCGTGTGCTGATTCATCAGTGGTGAATTATACCTTTGACCGAATGTGTTATCCAG 47296

QY      121 LeuLeuAsnGluCyAsnArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||||
DB      47297 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCCACCTCACTGCCAAG 47356

QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
      |||||
DB      47357 TCACATGAGACGGGTATATATGTTGATGATTTTTCAGATTGTGAATTTTGGCTGCC 47416

QY      161 LeuTyraAsnProPheGlyAsnPhelySProHisLeuGlnLySLeuCyAspGlyIleAsn 180
      |||||
DB      47417 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGTATCAAC 47476

QY      181 LysMetLeuAspGluGluAsnIle 188
      |||||
DB      47477 AAAATGTTGATGAAGAAACATA 47500

RESULT 8
US-10-097-065-98
; Sequence 98, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/097,065
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18

```

```

; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,008
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-097-065-98

Alignment Scores:
Pred. No.: 2,84e-96 Length: 1943
Score: 948.00 Matches: 186
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 2
Query Match: 98.54% Indels: 0
DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-097-065-98 (1-1943)
QY 1 MetAlaThraSpValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
DB 120 ATGGCCACGATGCTTTAATTCCAAAACCTGGCGTTCANGCACAAAGAAATCTTG 179
QY 21 GlyLysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 180 GGTAAATGCTGTCCAATCCATCGCCACCACTTAATGACACCAAGTGTGAGG 239
QY 41 LeuAspGluLeuTyrrArgValThrArgGluTyrrThrGlnAsnIysIysGluAlaGluLys 60
DB 240 CTGGATGAGCTCTACAGAGTGACACAGGAGTACACCAAAACAAGAGGAGGAGAG 299
QY 61 LysIleIysAsnLeuIleIysThrValIleIysIleAlaIleLeuTyrrArgAsnGln 80
DB 300 ATCATCAAAACCTCATCAAGACAGCATCAAGCTGGCCATCTTTATGGAATATCAG 359
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIleIysValIleGlnLeuAla 100
DB 360 TTTATCAAGATGAGTACATTGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT 419
QY 101 MetThrValIleSerPheIleGlnValAspTyrrThrPheAspArgAsnValIleSerArg 120
|||||
```

```

DB 420 ATGACCGTGTGATTCATCATCAGGTGATTAATCACTTTCAGCCGGAATGTATCCAGG 479
QY 121 LeuIleAsnGluCysArgGluMetLeuIleGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 480 CTGTAAATGAATGACAGAGATGCTGCACCAAAATCATTCACCGCCACTCATCTCCAG 539
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
DB 540 TCACATGACGGGTTAATATATGTGTTGATCATTTTCAGATTGTAATTTTGGCTGCC 599
QY 161 LeuTyrrAsnProPheGlyAsnIleLysProHisLeuGlnIysLeuCysAspGlyIleAsn 180
DB 600 TTGTATTAATCCTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGTCAAC 659
QY 181 LysMetLeuAspGluGlnAsnIle 188
DB 660 AAAATGTGTGATGAAGAAGAACTA 683

RESULT 9
US-10-372-876-98
; Sequence 98, Application US/10372876
; Publication No. US20030204071A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/372,876
; CURRENT FILING DATE: 2003-02-26
; PRIOR APPLICATION NUMBER: 09/334,595
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-372-876-98
```


Alignment Scores:

Pred. No.: 2,846-96 Length: 1943
Score: 948.00 Matches: 186
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 2
Query Match: 98.54% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-372-876-98 (1-1943)

```
QY 1 MetAlaThraSpVa1PheAsnSerLysAsnLeuAlaValGlnAglInLysLysIleLeu 20
DB 120 ATGGCCACAGATGCTTAAATTCAAAACCTGGCCCTTANGCACAAAAGAAATCTTG 179
QY 21 GlyLysMetValSerLysSerIleAlaThrLeuIleAspAspThrSerSerGluVal 40
DB 180 GGTAATAATGGTGTCCAAATCCATCCAGCCACCACTTAATAGCAGACAAAGATGAGAGTG 239
QY 41 LeuAspGluLeuTyraGValThrArgGluTyThrGlnAsnLysLysGluAglLys 60
DB 240 CTGGATGAGCTCTTACAGATGACCGAGGAGTACCCAAAACAAAGAGAGGAGAGAG 299
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraGAsnGln 80
DB 300 ATCATCAAGAACCTCATCAAGACGATCATCAGCTGGCCATCTTTATGAAATATCAG 359
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 360 TTTAATCAAGATGAGTACGATGATGAGAAATTTAAGAAAGAAAGTTCACTGAGCTTCT 419
QY 101 MetThrValAlaSerPheHisGlnValAspTyThrPheAspArgAsnValIleSerArg 120
DB 420 ATGACCGTGTGATGTTTCCATCAGTGTGATTAATCTTTAACCAGAAATGTTATTCAGG 479
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleGlnArgHisLeuThrAlaLys 140
DB 480 CTGTAAATGAATGACGAGAGAGATGCTGCACCAATCTTCAAGCCCACTCATCGCCAG 539
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 540 TCACATGACGAGCGGTATTAATATGTTGATCATTTTCAATTTGTAATTTTGGCTGCC 599
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 600 TTGTATATATCTTTTGGAAATTTTAAACCCCACTTACAAAACATGATGATGATCAAC 659
QY 181 LysMetLeuAspGluGlnAsnIle 188
DB 660 AAAATGTTGATGATGAAGAAACATA 683
```

RESULT 10

US-10-087-192-494
; Sequence 494, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 494
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-494

Alignment Scores:

Pred. No.: 4,576-93 Length: 2087
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 13 Gaps: 0

US-10-627-571-2 (1-188) x US-10-087-192-494 (1-2087)

```
QY 1 MetAlaThraSpVa1PheAsnSerLysAsnLeuAlaValGlnAglInLysLysIleLeu 20
DB 507 GTGGCTACAGATGCTTCAATTCAAAACCTGGCCCTTACAGGCACAAAAGAAATCTTG 566
QY 21 GlyLysMetValSerLysSerIleAlaThrLeuIleAspAspThrSerSerGluVal 40
DB 567 GGCAATAATGGTATCCAAATCCATCCAGCCACCACTGATGACGACACCAAGAGAGAGTG 626
QY 41 LeuAspGluLeuTyraGValThrArgGluTyThrGlnAsnLysLysGluAglLys 60
DB 627 CTAGATGAGCTCTTACAGGAGTACCAAGAGATGACCCCAAGAAAGAGAGAGAGAG 686
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraGAsnGln 80
DB 687 GTCATCAAGAACCTCATCAAGACGATCATCAGCTGGCCATCTTCCACAGAAACAACTCAG 746
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 747 TTCAATCAAGACGATGAGGCTCATGAGAAATTTCAAGAAAGAGAGAGAGAGAGAGAG 806
QY 101 MetThrValAlaSerPheHisGlnValAspTyThrPheAspArgAsnValIleSerArg 120
DB 807 ATGACCGTGTGATGTTTCCATCAGTGTGATTAATCTTTAACCAGAAATGTTATTCAGG 866
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleGlnArgHisLeuThrAlaLys 140
DB 867 CTGCTGAACGAGGCGGAGAGATGCTTCAACGAGATCATTTCAAGCCCACTTACCGCCAG 926
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 927 TCTCAGGAGCGGTATTAATATGCTTTGACCAATTTTCAATTTTGTATTTTGGCTGCC 986
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 987 TTGTACAAATCCCTTTGGAAGTTTAAACCTCACTTACAGAACTTTGCGAGCATCAAC 1046
QY 181 LysMetLeuAspGluGlnAsnIle 188
DB 1047 AAAATGTTGATGATGAAGAAACATA 1070
```

RESULT 11

US-10-087-192-493
; Sequence 493, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 62231
; TYPE: DNA
; ORGANISM: Mus musculus

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(62231)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-493

Alignment Scores:
Pred. No.: 7,56e-91 Length: 62231
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x US-10-087-192-493 (1-62231)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlylsIleLeu 20
   ::::
DB 50651 GTGGCTACAGATGCTTTCATTCCTCAAAACCTGGCGTTTCAGGACCAAAAGAGATCTCG 50710

QY 21 GlyIysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
   ::::
DB 50711 GGCAAAATGGTATCCAAATCCATGCGCACCGCTGATGACGACACCGACGAGGCTG 50770

QY 41 LeuAspGluLeuIlyrArgValThrArgIlyrThrGlnAsnIlysgluAlaGluys 60
   ::::
DB 50771 CTAGATGAGCTGTACAGGGGTGACCAAGAGATACACCCAGAACAAAGAGGCGGAGAG 50830

QY 61 IysIleIysAsnLeuIleIystrValIleIysLeuAlaIleLeuIlyrArgAsnAngin 80
   ::::
DB 50831 GTCATCAAAACCTCTCATAGAGCGGTATCAGCTGCGCTCTCCACAGGAACAATCAG 50890

QY 81 PheAsnGluAspGluLeuAlaLeuMetGluIysPheIlyIysValIhIsglnLeuAla 100
   ::::
DB 50891 TTCATCAAGACAGAGCTGGCGCTCATGAGAGATTCAGAGAGAGGACACAGCTTGCC 50950

QY 101 MetThrValIysSerPhehIsglnValAspIlyrThrPheAspArgAsnValLeuSerArg 120
   ::::
DB 50951 AAGAGCGTGTGAGCTTCCACCGAGTACAGTACCTTGAGCCGAGATGTGCTGCAG 51010

QY 121 LeuLeuAsnGluCysArgIleuMetLeuHISGlnIleIleGlnArgHISLeuThrAlaIys 140
   ::::
DB 51011 CTGGCTGAACGAGGCGGAGGCTCTTACACGAGATATTTCAGCGCACCTTACCGCAAG 51070

QY 141 SerHisGluYArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIys 160
   ::::
DB 51071 TCTCAGCGAGCGGTTTAATATGCTTTGACCATTTTTCAGATGTGATTTTGGCTGCC 51130

QY 161 LeuIlyrAsnProPheGluYAsnPhelysProHisLeuGlnIlyLeuIlyLeuIlyLeuAsn 180
   ::::
DB 51131 TTGTACATCCCTTTGGAAAGTTTAACCTTACAGAACTTTGGCAGCGGCATCAAC 51190

QY 181 LysMetLeuAspGluGluAsnIle 188
   ::::
DB 51191 AAAATGTGATGAAGAGAACATA 51214

RESULT 12
US-10-029-386-22914
; Sequence 22914, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22914
; LENGTH: 544

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; MAP TO AC035144.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EST_HUMAN HIT: AL549492.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P31390, EVALUE 2.60e-01
; OTHER INFORMATION: NT HIT: G17657123, EVALUE 0.00e+00
US-10-029-386-22914

Alignment Scores:
Pred. No.: 2.87e-93 Length: 544
Score: 914.00 Matches: 179
Percent Similarity: 99.44% Conservative: 0
Best Local Similarity: 99.44% Mismatches: 1
Query Match: 95.01% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x US-10-029-386-22914 (1-544)

QY 2 AlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlylsIleLeuGly 21
   ::::
DB 3 GGCACGATGCTTTTATTTCCAAAACCTGGCGTTTCAGGACCAAAAGAGATCTTGGGT 62

QY 22 LysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
   ::::
DB 63 AAAATGTGTCCAAATCCATCCACCAACCTTAAATAGACGACACAAAGTATGAGGTGCTG 122

QY 42 AspGluLeuIlyrArgValThrArgIlyrThrGlnAsnIlysgluAlaGluIlyls 61
   ::::
DB 123 GATGAGCTTACAGAGTGACGAGGATGACCCCAAAACAAAGAGAGGACAGAGATGC 182

QY 62 IleIysAsnLeuIleIystrValIleIysLeuAlaIleLeuIlyrArgAsnAnginPhe 81
   ::::
DB 183 ATCAAGAACCTCATCAAGCACTCATCAAGCTGCGCATTTTATAGGAATATCAGTTT 242

QY 82 AsnGluAspGluLeuAlaLeuMetGluIysPheIlyIysValIhIsglnLeuAlaMet 101
   ::::
DB 244 AATCAAGATGAGCTGATGATGAGAAATTAAGAAAGAAATTCATCAGCTTGCTATG 302

QY 102 ThrValIysSerPhehIsglnValAspIlyrThrPheAspArgAsnValLeuSerArg 121
   ::::
DB 303 ACCGCTGTGATTTCCATCAGGTGATGATTAACCTTTGACCGGAATGTATTCAGGCTG 362

QY 122 LeuAsnGluCysArgIleuMetLeuHISGlnIleIleGlnArgHISLeuThrAlaIysSer 141
   ::::
DB 363 TTAATGAAATGACAGAGATGCTGACCAAAATCATTCAGCGGCACCTCAGTCCAAAGTCA 422

QY 142 HisGluYArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleu 161
   ::::
DB 423 CATGACGGGTTTAATATGCTTTTATGATCATTTTTCAGATGTGAAATTTTGGCTGCC 482

QY 162 TyrAsnProPheGluYAsnPhelysProHisLeuGlnIlyLeuIlyLeuIlyLeuAsn 181
   ::::
DB 483 TATTAATCCTTTTGGGAATTTTAACCCCATTAACAAACATATGATGTATCAACAAA 542

RESULT 13
US-10-029-386-25411
; Sequence 25411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT FILING DATE: 2001-12-20

```

```

; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25411
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012678.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.51
; OTHER INFORMATION: SWISSPROT HIT: Q61768, EVALUATE 1.10e-01
; OTHER INFORMATION: NT HIT: AF120995.1, EVALUATE 1.30e-01
; OTHER INFORMATION: EST_HUMAN HIT: BF744133.1, EVALUATE 2.00e-59
US-10-029-386-25411

Alignment Scores:
Pred. No.: 5,49e-59 Length: 527
Score: 610.00 Matches: 106
Percent Similarity: 88.37% Conservative: 46
Best Local Similarity: 61.63% Mismatches: 20
Query Match: 63.41% Indels: 0
DB: 16 Gaps: 0

US-10-627-571-2 (1-188) x US-10-029-386-25411 (1-527)

QY 15 AlaGlnlyserlyleuGlylyserMetValSerlyserIleAlaThrThrIleuAlaP 34
DB 3 GCCAGAGAGAGATTCGAGCAAAATAGCCAGCAAACTGTGCCAATGTTGATTGAT 62
QY 35 AspThrserSerGluValLeuAspGluLeuTyraGlyValThrArgGluTyrrGlnAsn 54
DB 63 GACACCGACGACGAGATCTTGATGAGCTTACAAATCTCAAGAGCCACACACACAC 122
QY 55 LysLysGluAlaGluTyraGlylyserLysAsnIleuIleTyrrValIleLysLeuAlaIle 74
DB 123 AAGAGAGAGAGCCACCAAGATCATGAAGACTTATATCAAGTGGGAGATCAAAATCGGATC 182
QY 75 LeuTyrrArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGluTyrrPheLys 94
DB 183 CTCCTACCGGAACACCGATTAGCCAGAGAGCTGTTATTGTGGAGAGTCCGGAG 242
QY 95 LysValHisGlnLeuAlaMetThrValSerPheHisGlnValAspTyrrThrPheAsp 114
DB 243 AACCTGACCAACAGCCCATGACCATTTGTCAGCTTCTATGAGGGATACACCTTGAT 302
QY 115 ArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIleGln 134
DB 303 AGAGAGCTGCTCTCCATCTCCGTCATGAGTGAAGACCTGGTGCATGAACGTGTGAC 362
QY 135 ArgHisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSerAsp 154
DB 363 CGGCACCTGACGCGCCAGGACCCACGCGCGCATCAACACGCTTTAACTTTGCGCAT 422
QY 155 CysGluPheLeuAlaIleLeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLys 174
DB 423 GTGAGTTCCTCTCCACCTCTATAGCTGATGAGAGACGTGAGGCCCACTCAAGAG 482
QY 175 LeuCysAspGlyIleAsnLysMetLeuAspGluGlu 186
DB 483 ATTGTGAGAGATCAATTAAGTGTGATGAGAGAA 518

RESULT 14
US-10-094-466-43
; Sequence 43, Application US/10094466
; Publication No. US2003020363A1
; GENERAL INFORMATION:
; APPLICANT: Syntek et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281

```

```

; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 43
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(566)
US-10-094-466-43

Alignment Scores:
Pred. No.: 5.79e-57 Length: 619
Score: 593.00 Matches: 106
Percent Similarity: 79.14% Conservative: 42
Best Local Similarity: 56.68% Mismatches: 39
Query Match: 61.64% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-094-466-43 (1-619)

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGly 21
DB 5 GCCATGACACACTTCACGACCAAGAGCTGCTCTCAGCGGAGAGAGCTCTGAGT 64
QY 22 LysMetValSerlyserIleAlaThrThrIleuAlaAspAspThrserSerGluValLeu 41
DB 65 AAGATGCGCTCAAGGACAGTGTGCGCGCTGTGTGATGACACACAGCTGAGTCTG 124
QY 42 AspGluLeuTyraGlyValThrArgGluTyrrGlnAsnLysGluAlaGluTyrrLys 61
DB 125 GATGACGTGATCCGCGCCACAGGAGGTTCAAGCGGACGCGGAGGAGGCGGAGATG 184
QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnGlnPhe 81
DB 185 CTCAGAACCTGTCTCAAGGTGCGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTG 244
QY 82 AsnGluAspGluLeuAlaLeuMetGluTyrrPheLysLysValHisGlnLeuAlaMet 101
DB 245 GCGGTGAGAGAGTGTGCGCTGCTGCGCGCTTCCGACCGGCGCGCTGCGCATG 304
QY 102 ThrValAlaSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArgLeu 121
DB 305 ACGGCGGTGAGCTTCCACGAGGTGAGCTTCACTTCAAGCGGCGGCGGCGGCGG 364
QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleGlnArgHisLeuThrAlaLysSer 141
DB 365 CTGCTGAGTGGCGGACCTGCTGACAGGCGCGGTGCTCCACCTGAGCGGCAAGTCC 424
QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
DB 425 CACGGCGCATCAACACGTTTCCGCGACCTGAGCGGAGCTGCGGACTTCTGCTGCGGCTC 484

```


APPLICANT: Goodrich, Ryle
 APPLICANT: Aesund, Vinod
 APPLICANT: Ren, Feiyan
 APPLICANT: Xue, Aiding J.
 APPLICANT: Ma, Yunding
 APPLICANT: Wang, Zhiwei
 APPLICANT: Zhao, Qing A.
 APPLICANT: Zhang, Jie
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Dmanac, Radoje T.
 TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 791CIP2E
 CURRENT APPLICATION NUMBER: US/09/816,828
 CURRENT FILING DATE: 2001-03-22
 PRIOR APPLICATION NUMBER: 09/770,160
 PRIOR FILING DATE: 2001-01-26
 PRIOR APPLICATION NUMBER: 09/552,929
 PRIOR FILING DATE: 2000-04-18
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: pc_FL_genes Version 2.0
 SEQ ID NO 5
 LENGTH: 1188
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (174)..(731)
 US-09-816-828-5
 Alignment Scores:
 Pred. No.: 2,6e-56 Length: 1188
 Score: 591.00 Matches: 106
 Percent Similarity: 79.14% Conservative: 42
 Best Local Similarity: 56.68% Mismatches: 39
 Query Match: 61.43% Indels: 0
 DB: Gaps: 0
 US-10-627-571-2 (1-188) x US-09-816-828-5 (1-1188)
 QY 2 AlTThAspValPheAmSerlyAsnLeuAlaValGlnAlGlnlyblyslleuGly 21
 Db 171 GCGATGACACCTTCAGCCACCAAGAGCCTGCTCGACGGGCGAGAAAGACCTCTAGT 23
 QY 22 LyMeValSerlyserlleAlaThrThrleuileAspAspTrnSerSerGluValleu 41
 Db 231 AAGATGCGCTCAAGGACAGTGTGGCCCTGCTGGTGATGACACCGAGAGTGAAGTCTCTG 290
 QY 42 AspGluLeuTyArgValThzArgLutTyThzGlnAenlyblyGluAlaGluTylys 61
 Db 291 GATGACTGTACCGGCCACCGAGAGTTCACCGCCAGCCCGCAGAGGCCCAAGATG 350
 QY 62 lIeLyAsnleuilelyeThrValilelyslleuAlaileuTyArgAsnAmGlnPhe 81
 Db 351 CTCAGAACTGGTCAAGGTGAGCTGGAAGCTGGAGTCTGCTGCGTGGGAGACACCTG 410
 QY 82 AsnGlnAspGluLeuAlaLeuMetGluyspelyslyblyValIhIsglnleuAlaMet 101
 Db 411 GGGCGTGAAGACTGGCGCTGCTGGGGGCTTCCGCCACCGGGCGCGCTGCTGGCCATG 470
 QY 102 ThValValSerPhehIsglnValAspTyTrhPheAspArgAmValleuSerArgLeu 121
 Db 471 ACGGCGCTAGCTTCACCAAGTGAATTCACCTTCACCGGCGCGTGTGGCGGG 530
 QY 122 LeuAsnGluCyArgGluMetLeuHhIsglnIleleGlnArgHIsleuThralaIysSer 141
 Db 531 CTGCTCGATGTCGCGGACCTGCTGACACAGGCGCGGTGCTCCACCTGACCGCAATGCC 590
 QY 142 HIsGlyArgValAsnAsnValPheAspHisPheSerAspCyArgLupheleuAlaIleu 161
 Db 591 CACGGCGGATCAACCAAGTGTGGCCACCTTACGCCACCTGCGGCTGCGCGCTC 650
 QY 162 TyAsnProPheGlyAsnPheIysProHIsleuGlnIlyslleuCyAspGlyIleAsnIys 181

```
Db      651 TACGGCCCGCCGAGCCCTACCGCTCCACCTCGCAGAGATGTGCGAGGGCGCTGGGCGG 710
QY      182 MetLeuAspGluGluAsnIle 188
       |||||
Db      711 ATGCTGACGAGGGCGAGCCTC 731

RESULT 18
US-10-764-420-2116
; Sequence 2116, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lam, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methode For Determining Whether An Agent
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 80
; OTHER INFORMATION: n = A,T,C or G
US-10-764-420-2116

Alignment Scores:
Pred. No.:      2,18e-53      Length:      1389
Score:          566.00      Matches:      101
Percent Similarity: 79.46%      Conservative: 46
Best Local Similarity: 54.59%      Mismatches: 38
Query Match:    58.84%      Indels:      0
DB:            21      Gaps:      0

US-10-627-571-2 (1-188) x US-10-764-420-2116 (1-1389)
QY      4 AspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIleLeuGlyIleLeuGlyIleValMet 23
       |||||
Db      87 GACACCTTCAGCAGGAAGCCTGGCCCTGCAGGCCAGGAAGTCTCTCAGCAAGATG 146
QY      24 ValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValIleuAspGlu 43
       |||||
Db      147 GCCTCCAAAGCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 206
QY      44 LeuTyrArgValThrArgGluTyrThrGlnAsnIleValGlnIleValGlnIleValGlnIleVal 63
       |||||
Db      207 CTGTACCAAGCCAGCAGAGGATTCACGCGCAGCGAGAGGACACAGAGGATGTAAGTGAAG 266
QY      64 AsnLeuIleIleThrValIleIleValLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGln 83
       |||||
Db      267 AACCTGTGTAAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 326
QY      84 AspGluLeuAlaLeuMetGluIysPheIysValIleGlnLeuAlaMetThrVal 103
       |||||
Db      327 AATGAGCTGGCCCAAGCTGACGCGGTTCCGGGGCCGGGTCCGCAACTGGCCATGACAGCC 386
QY      104 ValSerPheHisGlnValAspTyrThrPheAspArgAsnValIleuSerArgLeuLeuAsn 123
       |||||
Db      387 CTCAGCTTCACACAGGATGACTTCACTTGTACCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 446
QY      124 GluGlyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIysSerHisGly 143
       |||||
Db      447 GAGTGACAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 506
```

```
QY      144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTyrAsn 163
       |||||
Db      507 CCATCATCATCATCTTCTAGTCACTTTGGCAATGTGTACTTCTGGCCGCGCTGTACAC 566
QY      164 PropheGlyAsnPheIysProHisLeuGlnIleValCysAspGlyTyrIleAsnIysMetLeu 183
       |||||
Db      567 CCAGCAGAGCCCTACCGGAGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 626
QY      184 AspGluGluAsnIle 188
       |||||
Db      627 GACGAGGTGTGCATC 641

RESULT 19
US-09-822-849A-81
; Sequence 81, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 81
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-81

Alignment Scores:
Pred. No.:      3,92e-49      Length:      1113
Score:          527.00      Matches:      99
Percent Similarity: 77.22%      Conservative: 40
Best Local Similarity: 55.00%      Mismatches: 39
Query Match:    54.78%      Indels:      2
DB:            9      Gaps:      1

US-10-627-571-2 (1-188) x US-09-822-849A-81 (1-1113)
QY      6 PheAsnSerIysAsnLeuAlaValGlnIleValIleLeuGlyIleValMetValSer 25
       |||||
Db      72 TTCAGCTCAAAAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 131
QY      26 LysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValIleuAspGluLeuTyr 45
       |||||
Db      132 CCTCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 191
QY      46 ArgValThrArgGluTyrThrGlnAsnIleValGlnIleValGlnIleValGlnIleValGlnIle 65
       |||||
Db      192 CGTGTGTCAAGAGATGACAGCAGCGCGGCCCGCCAGCGCGGTATCAAGACCTG 251
QY      66 IleValThrValIleIleValLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnIleAspGlu 85
       |||||
Db      252 ATCAAAAGTGGCCATCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 311
QY      86 LeuAlaLeuMetGluIysPheIysValIleGlnLeuAlaMetThrValValSer 105
       |||||
Db      312 CTGGCCCTGTGTACCCGCTTTTGTCCAGAGAGCTGCGGACGGGTGCAATGACGACCTTAC 371
QY      106 PheHisGlnValAspTyrThrPheAspArgAsnValIleuSerArgLeuLeuAsnGluCys 125
       |||||
```

Db 372 TTGTGATGAGTACCTTACCTTCGAGCGTCTGCTGGCTGGCTGACCGAGTGC 431
 QY 126 ArgGlMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIysSerHisGlyArgVal 145
 Db 432 CGGATGTGCTGTAGAGTTGGAGAACACCACTACGCCCAAGTCACTAGCGCCATC 491
 QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTyraAsnProPhe 165
 Db 492 CGCCACGTGTTTATCATCTTCTCTGACCCAGGTCCTGACGCCCTCTATGGGCTT 548
 QY 166 GlyAsnPhelYsPheHisLeuGlnIleGlnCysAspGlyIleAsnIysMetLeuAspGlu 185
 Db 549 ---GACTTCACTACGACCACTTGGCAAGATCTGTGACGAGTCAAGAAAGCTGTAAGCA 605

RESULT 20
 US-10-363-374-15
 ; Sequence 15, Application US/10363374
 ; Publication No. US20050048483A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Su, Eric
 ; APPLICANT: Wang, He
 ; TITLE OF INVENTION: NOVEL SECRETED PROTEINS AND THEIR USES
 ; FILE REFERENCE: X13020
 ; CURRENT APPLICATION NUMBER: US/10/363,374
 ; CURRENT FILING DATE: 2003-02-27
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 15
 ; LENGTH: 1165
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: CDS
 ; LOCATION: (100)..(651)
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: (100)..(180)
 US-10-363-374-15

Alignment Scores:
 Pred. No.: 4,26-49 Length: 1165
 Score: 527.00 Matches: 99
 Percent Similarity: 77.22% Conservative: 40
 Best Local Similarity: 55.00% Mismatches: 39
 Query Match: 54.78% Indels: 2
 Gaps: 1

US-10-627-571-2 (1-188) x US-10-363-374-15 (1-1165)

QY 6 PheAsnSerIysAsnLeuAlaValGlnIleGlnIleLeuGlyIysMetValSer 25
 Db 109 TTGAGCTCAAAAGAGCTGCGACCTGCAAGCAGAGAGAGAGTACTAGTAGAGCGGGT 168
 QY 26 LysSerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluTyr 45
 Db 169 CGCTGTGTGCTCATCTCTTCATAGAGTACAGCAAGAGTGTAGTGTAGTGTAGTCTAC 228
 QY 46 ArgValThrArgGluTyrThrGlnAsnIleGlnIleGlnIleGlnIleGlnIleGlnIle 65
 Db 229 CGTGTGTCCCAAGAGTACACGACAGCGCGGCCCGCCAGCGCGGTATCAAGACCTG 288
 QY 66 IleValThrValIleIleIleLeuAlaIleLeuTyraArgAsnAsnIlePheAsnGlnAspGlu 85
 Db 289 ATCAAAAGTGCATCAAGAGT 348
 QY 86 LeuAlaLeuMetGluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 105
 Db 349 CTGGCCCTGCTACCGCGCTTCCGCAAGCTGCGGCGGAGGTGTCATGACGACGACCTTAGC 408
 QY 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125
 Db 409 TTGTGTGAGGTAGACCTTCACTTCAGAGCGTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 468

QY 126 ArgGlMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIysSerHisGlyArgVal 145
 Db 469 CGGATGTGCTGTAGAGTTGGAGAACACCACTACGCCCAAGTCACTAGCGCCATC 528
 QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTyraAsnProPhe 165
 Db 529 CGCCACGTGTTTATCATCTTCTCTGACCCAGGTCCTGACGCCCTCTATGGGCTT 585
 QY 166 GlyAsnPhelYsPheHisLeuGlnIleGlnCysAspGlyIleAsnIysMetLeuAspGlu 185
 Db 586 ---GACTTCACTACGACCACTTGGCAAGATCTGTGACGAGTCAAGAAAGCTGTAAGCA 642

RESULT 21
 US-10-302-172-537
 ; Sequence 337, Application US/10302172
 ; Publication No. US20040053250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Demanac, Radoje T.
 ; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids a
 ; FILE REFERENCE: 803 1CNP
 ; CURRENT APPLICATION NUMBER: US/10/302,172
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/225,251
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: PCT US02/05095
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 09/799,451
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 950
 ; SOFTWARE: pt_genes Version 2.0
 ; SEQ ID NO 537
 ; LENGTH: 1175
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: CDS
 ; LOCATION: (113)..(664)
 US-10-302-172-537

Alignment Scores:
 Pred. No.: 4,26-49 Length: 1175
 Score: 527.00 Matches: 99
 Percent Similarity: 77.22% Conservative: 40
 Best Local Similarity: 55.00% Mismatches: 39
 Query Match: 54.78% Indels: 2
 Gaps: 1

US-10-627-571-2 (1-188) x US-10-302-172-537 (1-1175)

QY 6 PheAsnSerIysAsnLeuAlaValGlnIleGlnIleLeuGlyIysMetValSer 25
 Db 122 TTGAGCTCAAAAGAGCTGCGACCTGCAAGCAGAGAGAGTACTAGTAGAGCGGGT 181
 QY 26 LysSerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluTyr 45
 Db 182 CGCTGTGTGCTCATCTCTTCATAGTACAGAGAGTGTAGTGTAGTGTAGTGTAGTCTAC 241
 QY 46 ArgValThrArgGluTyrThrGlnAsnIleGlnIleGlnIleGlnIleGlnIleGlnIle 65
 Db 242 CGTGTGTCCCAAGAGTACACGACAGCGCGGCCCGCCAGCGCGGTATCAAGACCTG 301
 QY 66 IleValThrValIleIleIleLeuAlaIleLeuTyraArgAsnAsnIlePheAsnGlnAspGlu 85
 Db 302 ATCAAAAGTGCATCAAGAGT 361
 QY 86 LeuAlaLeuMetGluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 105
 Db 362 CTGGCCCTGCTACCGCGCTTCCGCAAGCTGCGGCGGAGGTGTCATGACGACGACCTTAGC 421
 QY 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125

```
DB 422 TTGGTGAAGTGAAGTCACTTCGAGGCTGCTGTTCTGGCTGCGCTGACCGAGTGC 481
QY 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyVal 145
DB 482 CGGAGTGTGCTGTAAAGTGGGAAACACACTTACCGCCCAAGTCACTGGCCGCT 541
QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTrpAsnProPhe 165
DB 542 CGCAGAGTGTGATCACTTCTGTGACCGAGCTGTGCTACGCGCTCTATGGGCT--- 598
QY 166 GlyAsnPheLysProHisLeuGlnIlyLeuCysAspGlyIleAsnLysMetLeuAspGlu 185
DB 599 ---GACTTCACTCAAGCACCCTTGGCAAGATCTGTGACGAGCTCAAGAGCTGTAAGCA 655

RESULT 22
US-10-959-539-59
; Sequence 59, Application US/10959539
; Publication No. US20050048623A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BANDMAN, Olga
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YANG, Junming
; APPLICANT: LU, Dying Aina M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: SHAH, Putvi
; TITLE OF INVENTION: CELL CYCLE AND PROLIFERATION PROTEINS
; FILE REFERENCE: PF-0722 PCT
; CURRENT APPLICATION NUMBER: US/10/959,539
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/10/031,915
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/145,075; 60/153,129; 60/164,647
; PRIOR FILING DATE: 1999-07-21; 1999-09-08; 1999-11-10
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PERL Program
; SEQ ID NO 59
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1558289CB1
US-10-959-539-59

Alignment Scores:
Pred. No.: 4,77e-49 Length: 1268
Score: 527.00 Matches: 99
Percent Similarity: 77.22% Conservative: 40
Best Local Similarity: 55.00% Mismatches: 39
Query Match: 54.78% Indels: 2
DB: 21 Gaps: 1

US-10-627-571-2 (1-188) x US-10-959-539-59 (1-1268)
QY 6 PheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlnLysMetValSer 25
DB 112 TTCAGTCAAAAGAGCTTGGCACTGCAAGCAGAGAAAGAGTACTGATGAATGAGCGGCT 171
QY 26 LysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTrp 45
DB 172 CGCTGTGTGCTCATCTCTTCATAGATGAGACAAGCAGAGTGTGCTAGTATGACTCTAC 231
QY 46 ArgValThrArgGluIlyThrGlnAsnLysLysGluAlaGlnLysLysIleLysAsnLeu 65
DB 232 CGTGTGTCAAGAGTACAGCAGCAGCGCGCCAGCGCCAGCGTGTATCAAGGACTGTG 291
```

```
QY 66 IleLysThrValIleLysLeuAlaIleLeuTrpArgAsnGlnPheAsnGlnAspGlu 85
DB 292 ATCAAAAGTGCATCAAGGTGTGCTGTGCTGACCGCCAAATGCTCTTGGCCCACTGAG 351
QY 86 LeuAlaLeuMetGlnLysPheLysLysValHisGlnLeuAlaMetThrValValSer 105
DB 352 CTGGCCCTGTGTAACCGCTTGTGCGAAGAGTGGCGGAGGGTGCATGACCGCACTTAGC 411
QY 106 PheHisGlnValAspTrpThrPheAspArgAsnValLeuSerArgLeuLeuAsnLys 125
DB 412 TTGGTGAAGTGAAGTCACTTCGAGGCTGCTGTTCTGGCTGCGCTGTGACCGAGTGC 471
QY 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyVal 145
DB 472 CGGAGTGTGCTGTGAGTGTGAGAAACACACTTCAAGCCCAAGTCACTGACCGGCTATC 531
QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTrpAsnProPhe 165
DB 532 CGCAGTGTGATCACTTCTGTGACCGAGCTGTGCTACGCGCTCTATGGGCT--- 588
QY 166 GlyAsnPheLysProHisLeuGlnIlyLeuCysAspGlyIleAsnLysMetLeuAspGlu 185
DB 589 ---GACTTCACTCAAGCACCCTTGGCAAGATCTGTGACGAGCTCAAGAGCTGTAAGCA 645

RESULT 23
US-09-796-692-3207
; Sequence 3207, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Manion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3207
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-3207

Alignment Scores:
Pred. No.: 2.01e-47 Length: 287
Score: 504.00 Matches: 95
```


Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.39% Indels: 0
 DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x US-09-796-692-3207 (1-287)

QY 74 ||leuetyrargananglnphasnnglnaspglnleu|aleuemetglnlysphe|ys 93
 |||||
 DB 3 ATTCTTAATGGAATTAATCAAGTTTAATCAAGATGAGCTTGAATGAGTAATGAGTAATTAAG 62
 QY 94 LysLysValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyrThrPhe 113
 |||||
 DB 63 AAGAAGTTCACTACGCTTGCTATACCGTGTCCATCCAGTGGATTAATACCTTT 122
 QY 114 AsparGsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIle 133
 |||||
 DB 123 GACCGAATGTGTTATCCAGGCTGTTAAATGACAGAGATGCTGCACCAATCATTT 182
 QY 134 GlnArgHisLeuThrAlaLysSerHisGlyArgValAsnValPheAspHisPheSer 153
 |||||
 DB 183 CAGGCCACCTCCTCAGCTGCACAGTCACATGACGGGTTAATATGTTGATCATTTTCA 242
 QY 154 AspCysGlnPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPhe 168
 |||||
 DB 243 GATTGGAATTTTGGCTGCTGTATATCTTTTGGGAATTTT 287

RESULT 24

US-10-040-862-3207
 ; Sequence 3207, Application US/10040862
 ; Publication No. US20030078396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Corixa Corporation
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
 ; FILE REFERENCE: 014058-013520US
 ; CURRENT APPLICATION NUMBER: US/10/040,862
 ; CURRENT FILING DATE: 2001-11-06
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: US 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: US 09/796,692
 ; PRIOR FILING DATE: 2001-03-01
 ; NUMBER OF SEQ ID NOS: 10467
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3207
 ; LENGTH: 287
 ; TYPE: DNA

ORGANISM: Homo sapiens
 US-10-040-862-3207

Alignment Scores:

Score: 2,01e-47 Length: 287
 504.00 Matches: 95
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.39% Indels: 0
 DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-040-862-3207 (1-287)

QY 74 ||leuetyrargananglnphasnnglnaspglnleu|aleuemetglnlysphe|ys 93
 |||||
 DB 3 ATTCTTAATGGAATTAATCAAGTTTAATCAAGATGAGCTTGAATGAGTAATTAAG 62
 QY 94 LysLysValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyrThrPhe 113
 |||||
 DB 63 AAGAAGTTCACTACGCTTGCTATACCGTGTCCATCCAGTGGATTAATACCTTT 122
 QY 114 AsparGsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIle 133
 |||||
 DB 123 GACCGAATGTGTTATCCAGGCTGTTAAATGACAGAGATGCTGCACCAATCATTT 182
 QY 134 GlnArgHisLeuThrAlaLysSerHisGlyArgValAsnValPheAspHisPheSer 153
 |||||
 DB 183 CAGGCCACCTCCTCAGCTGCACAGTCACATGACGGGTTAATATGTTGATCATTTTCA 242
 QY 154 AspCysGlnPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPhe 168
 |||||
 DB 243 GATTGGAATTTTGGCTGCTGTATATCTTTTGGGAATTTT 287

RESULT 25

US-10-057-475B-3207
 ; Sequence 3207, Application US/10057475B
 ; Publication No. US2004002068A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Clapper, Jonathan David
 ; APPLICANT: Wang, Aljun
 ; APPLICANT: Ordenez, Nadia
 ; APPLICANT: Carter, Lauren
 ; APPLICANT: McNeill, Patricia Dianne
 ; APPLICANT: Corixa Corporation
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
 ; FILE REFERENCE: 014058-014402US
 ; CURRENT APPLICATION NUMBER: US/10/057,475B
 ; CURRENT FILING DATE: 2002-01-22
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; Remaining Prior Application data removed - See File Wrapper or PALM.

```
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-3207
```

Alignment Scores:

Pred. No.:	2,01e-47	Length:	287
Score:	504.00	Matches:	95
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.39%	Indels:	0
DB:	17	Gaps:	0

US-10-627-571-2 (1-188) x US-10-057-475B-3207 (1-287)

```
QY 11leuTYrArGAsnAnngInPheAsngInAspGluLeuAlaLeuMeGluLysPheLys 93
DB 3 ATCTTTATAGAAATATACGTTTAATCAAGATGAGCTAGTCATGAGAAATTTAAG 62
QY 94 LysIysValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyrThrPhe 113
DB 63 AAGAAAGTTCATCAGCTGCTATGACCGGTGTCAGTTTCATCAGGTGATTAACCTTT 122
QY 114 AspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIleIle 133
DB 123 GACCGAAATGTGTTATCCAGGCTGTTAAATGAGACAGATGCTGCCAATAATCATTT 182
QY 134 GlnArgHisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSer 153
DB 183 CAGCGCACCTCCTCAGCAAGTCACATGACGCGGTTAATATGTTGATTCATTTTCA 242
QY 154 AspCysGluPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPhe 168
DB 243 GATGTGAATTTTGGCTGCTGTATATCCTTTGGGAATTTT 287
```

RESULT 26

```
US-10-154-884B-3207
/ Sequence 3207, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
```

/ Remaining Prior Application data removed - See File Wrapper or PALM.

```
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-154-884B-3207
```

Alignment Scores:

Pred. No.:	2,01e-47	Length:	287
Score:	504.00	Matches:	95
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.39%	Indels:	0
DB:	17	Gaps:	0

US-10-627-571-2 (1-188) x US-10-154-884B-3207 (1-287)

```
QY 74 11leuTYrArGAsnAnngInPheAsngInAspGluLeuAlaLeuMeGluLysPheLys 93
DB 3 ATCTTTATAGAAATATACGTTTAATCAAGATGAGCTAGTCATGAGAAATTTAAG 62
QY 94 LysIysValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyrThrPhe 113
DB 63 AAGAAAGTTCATCAGCTGCTATGACCGGTGTCAGTTTCATCAGGTGATTAACCTTT 122
QY 114 AspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIleIle 133
DB 123 GACCGAAATGTGTTATCCAGGCTGTTAAATGAGACAGATGCTGCCAATAATCATTT 182
QY 134 GlnArgHisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSer 153
DB 183 CAGCGCACCTCCTCAGCAAGTCACATGACGCGGTTAATATGTTGATTCATTTTCA 242
QY 154 AspCysGluPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPhe 168
DB 243 GATGTGAATTTTGGCTGCTGTATATCCTTTGGGAATTTT 287
```

RESULT 27

```
US-10-764-324-3207
/ Sequence 3207, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ CURRENT FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-08-03
```

```

; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3207
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-3207

Alignment Scores:
Pred. No.: 2,01e-47      Length: 287
Score: 504.00           Matches: 95
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 52.39%      Indels: 0
DB: 19                  Gaps: 0

US-10-627-571-2 (1-188) x US-10-764-324-3207 (1-287)

QY 74 ILEUTYARGAANANGINPHEANGINASPGLULEUALALEUWETGIULYSPHELYS 93
DB 3 ATCTTTATAGGATATATCATGTTTAATCAAGATGAGCTACATTTGATGAGAAATTTAAG 62
QY 94 LVELYVALHISGLINLEUALAMETTRVALISERPHHISGLINVALASPTYRTHRPH 113
DB 63 AAGAAAGTTATCATGCTTGCTATGACCGTGCTCATGTTTCATCATGAGTGATTAACCTTT 122
QY 114 ASPARGANVALLEUSERARGLEULEAENGLUCYARGSLUMETLEUHSGLINLEILE 133
DB 123 GACCGGATGTGTTATTCACGCTGTTAAATGATGACAGAGATGCTGCACCAATCATTT 182
QY 134 GLNARGHISLEUTHRALALYSSEHISGLYARGVALASNAVALPHEASPHISPHESER 153
DB 183 CACGGCCACCTCATCTGCCAAGTCACATGACGCGGTTAATGTTGATCATTTTTC 242
QY 154 ASPCYGSLUPHEULALALALEUTYRAENPROPHGLYASNPHE 168
DB 243 GATTGTGAATTTTGGCTGCTGTATATTCCTTTGGGAATTTT 287

RESULT 28
US-09-796-692-2801/c
; Sequence 2801, Application US/09796692
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
```

```

; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-2801

Alignment Scores:
Pred. No.: 7.37e-47      Length: 287
Score: 499.00           Matches: 94
Percent Similarity: 100.00%  Conservative: 1
Best Local Similarity: 98.95%  Mismatches: 0
Query Match: 51.87%      Indels: 0
DB: 9                  Gaps: 0

US-10-627-571-2 (1-188) x US-09-796-692-2801 (1-287)

QY 74 ILEUTYARGAANANGINPHEANGINASPGLULEUALALEUWETGIULYSPHELYS 93
DB 285 ATCTTTATAGGATATATCATGTTTAATCAAGATGAGCTACATTTGATGAGAAATTTAAG 226
QY 94 LVELYVALHISGLINLEUALAMETTRVALISERPHHISGLINVALASPTYRTHRPH 113
DB 225 AAGAAAGTTATCATGCTTGCTATGACCGTGCTCATGTTTCATCATGAGTGATTAACCTTT 166
QY 114 ASPARGANVALLEUSERARGLEULEAENGLUCYARGSLUMETLEUHSGLINLEILE 133
DB 165 GACCGGAAATGTGTTATTCACGCGCTGTTAAATGATGACAGAGATGCTGCACCAATCATTT 106
QY 134 GLNARGHISLEUTHRALALYSSEHISGLYARGVALASNAVALPHEASPHISPHESER 153
DB 105 CACGGCCACCTCATCTGCCAAGTCACATGACGCGGTTAATGATGTTTTCATTTTTC 46
QY 154 ASPCYGSLUPHEULALALALEUTYRAENPROPHGLYASNPHE 168
DB 45 GATTGTGAATTTTGGCTGCTGTATATTCCTTTGGGAATTTT 1

RESULT 29
US-10-040-862-2801/c
; Sequence 2801, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
```

```
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-2801

Alignment Scores:
Pred. No.: 7,37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-040-862-2801 (1-287)

QY 74 l l e l e u t y r a r g a s n a n g l i n p h e a n g l i n a s p g l u l e u a l a l e u m e t g l u l y s p h e l y s 93
DB 285 a t t c t t t a a g g a t a t a t c a g t t t a a t c a a g a t g a c t a t g a t g a g a a a t t t a a g 226
QY 94 l y s l y s v a l i h i s g l i n e u a l a m e t h r v a l s e r p h e i s g l i n v a l a s p y r t h r p h e 113
DB 225 a a g a a a g t t c a t a g c t t g t t a t a c c g g t c a g t t t c a t a c a g t g a t t a t a c c t t t 166
QY 114 a s p a r g a s n a l l e u s e r a r g l e u l e u a n g l u c y a r g l u m e t l e u h i s g l i n l e l l e 133
DB 165 g a c c g g a a t g t t a t t c c a g c c t g t t a a t g a a t g a g a n t g c t g c a c c a a t c a t t 106
QY 134 g l n a r g h i s l e u t h r a l a l y s e r h i s g l y a r g v a l a s n a s n v a l p h e a s p h i s p h e s e r 153
DB 105 c a g c g c c a c c t c a c t g c c a a g t c a c a t g a c g g g t t a a t a t g t t t g a t c a t t t t t t c a 46
QY 154 a a p c y s g l u p h e l e u a l a l a l e u t y r a s n p r o p h e g l y a s n p h e 168
DB 45 g a t t g t g a a t t t t t t g c t g c c t t g t a t a t c t t t t t g g a a t t t t 1

RESULT 30
US-10-057-475B-2801/c
; Sequence 2801, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01440205
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
```

```
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-2801

Alignment Scores:
Pred. No.: 7,37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-057-475B-2801 (1-287)

QY 74 l l e l e u t y r a r g a s n a n g l i n p h e a n g l i n a s p g l u l e u a l a l e u m e t g l u l y s p h e l y s 93
DB 285 a t t c t t t a a g g a t a t a t c a g t t t a a t c a a g a t g a c t a t g a t g a g a a a t t t a a g 226
QY 94 l y s l y s v a l i h i s g l i n e u a l a m e t h r v a l s e r p h e i s g l i n v a l a s p y r t h r p h e 113
DB 225 a a g a a a g t t c a t a g c t t g t t a t a c c g g t c a g t t t c a t a c a g t g a t t a t a c c t t t 166
QY 114 a s p a r g a s n a l l e u s e r a r g l e u l e u a n g l u c y a r g l u m e t l e u h i s g l i n l e l l e 133
DB 165 g a c c g g a a t g t t a t t c c a g c c t g t t a a t g a a t g a g a n t g c t g c a c c a a t c a t t 106
QY 134 g l n a r g h i s l e u t h r a l a l y s e r h i s g l y a r g v a l a s n a s n v a l p h e a s p h i s p h e s e r 153
DB 105 c a g c g c c a c c t c a c t g c c a a g t c a c a t g a c g g g t t a a t a t g t t t g a t c a t t t t t c a 46
QY 154 a a p c y s g l u p h e l e u a l a l a l e u t y r a s n p r o p h e g l y a s n p h e 168
DB 45 g a t t g t g a a t t t t t t g c t g c c t t g t a t a t c t t t t t g g a a t t t t 1

RESULT 31
US-10-154-884B-2801/c
; Sequence 2801, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01352105
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
```

```

; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-2801

Alignment Scores:
Pred. No.: 7.37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x US-10-154-884B-2801 (1-287)
QY 74 ILeuTyRrAgAnGnInPhaAnGlaSPgluLeuAlaLeuMetGluLysPheLys 93
DB 285 ATTCTTATAGGGATATACGTTTAAATCAAGATGAGCATGATGAGAAATTTAAG 226
QY 94 LysLysValHisGlnLeuAlaMetThrValLysSerPheHisGlnValAspTyrThrPhe 113
DB 225 AAGAAAGTTATCATCGCTTGTCATGACCGTGTCAAGTTCCATGACGTGGATTAACCTTT 166
QY 114 AspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIle 133
DB 165 GACCGGATGTGTTATCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCATTT 106
QY 134 GluArgHisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSer 153
DB 105 CAGCGCACCTCATCTGCACAGTCACATGACGCGGTTAATATGTTGATCATTTTCA 46
QY 154 AspCysGluPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPhe 168
DB 45 GATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 1

RESULT 32
US-10-764-324-2801/c
; Sequence 2801, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Hematological and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
```

```

; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-2801

Alignment Scores:
Pred. No.: 7.37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x US-10-764-324-2801 (1-287)
QY 74 ILeuTyRrAgAnGnInPhaAnGlaSPgluLeuAlaLeuMetGluLysPheLys 93
DB 285 ATTCTTATAGGGATATACGTTTAAATCAAGATGAGCATGATGAGAAATTTAAG 226
QY 94 LysLysValHisGlnLeuAlaMetThrValLysSerPheHisGlnValAspTyrThrPhe 113
DB 225 AAGAAAGTTATCATCGCTTGTCATGACCGTGTCAAGTTCCATGACGTGGATTAACCTTT 166
QY 114 AspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIle 133
DB 165 GACCGGATGTGTTATCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCATTT 106
QY 134 GluArgHisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSer 153
DB 105 CAGCGCACCTCATCTGCACAGTCACATGACGCGGTTAATATGTTGATCATTTTCA 46
QY 154 AspCysGluPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPhe 168
DB 45 GATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 1

RESULT 33
US-09-796-692-5724
; Sequence 5724, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
```


Qy 153 rAepCyGsluPheleuAlaIaleuTyraenProPheGlyasnPh 168
Db 243 AGATTGGAATTTTGGCTGCTGTATATCTTTGGGAATTTT 288

RESULT 35
US-10-057-475B-5724
; Sequence 5724, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Mang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-0140205
; CURRENT FILING DATE: 2002-01-22
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(288)
; OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-5724

Alignment Scores:
Pred. No.: 3.65e-45 Length: 288
Score: 484.00 Matches: 94
Percent Similarity: 97.92% Conservative: 0
Best Local Similarity: 97.92% Mismatches: 1
Query Match: 50.31% Indels: 1
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-057-475B-5724 (1-288)

Qy 74 lIeLeuTyraGsaBsaNgInPheAaNgInAspGInleuAlaIeueWetGluYs-PheLy 93
Db 3 ATTCTTATGATGATATATCAAGATGACCTAGCATGATGAGGAATNNCAA 62
Qy 93 sLYsVaiHIGInleuAlaIeueWetThrValIsePheHIGInValAspTYrThrPh 113

Db 63 GAAGAAAGTTTCATCAGCTTGCTATGACCGTGTCAGTTTCATCAGGTGATTAACCTT 122
Qy 113 eAaPaTgaBsaValIeueSerArgleuIeueaNgInuCyAaGsluMetLeuHIGInIleI 133
Db 123 TGACCCGAATGCTTATCCAGGCTGTTAAATGATGACAGATGCTGCACCAATCAT 182
Qy 133 eGInaRgHIGInleuThAlaIysSerHIGInYArGValaAsaBsaValPheAspHISpHe 153
Db 183 TCAGCCGACCTCAGTCCAGTCAAGTCAAGTCAAGGCTTAAATATGTGTTTATCTTTTC 242

Qy 153 rAepCyGsluPheleuAlaIaleuTyraenProPheGlyasnPh 168
Db 243 AGATTGGAATTTTGGCTGCTGTATATCTTTGGGAATTTT 288

RESULT 36
US-10-154-884B-5724
; Sequence 5724, Application US/1015484B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01352105
; CURRENT FILING DATE: 2002-05-23
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(288)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-5724

Alignment Scores:
Pred. No.: 3.65e-45 Length: 288
Score: 484.00 Matches: 94
Percent Similarity: 97.92% Conservative: 0
Best Local Similarity: 97.92% Mismatches: 1
Query Match: 50.31% Indels: 1
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-154-884B-5724 (1-288)

Qy 74 lIeLeuTyraGsaBsaNgInPheAaNgInAspGInleuAlaIeueWetGluYs-PheLy 93

```

Db      3  ATTCTTATAGAGATATATGATTATATCAAGATGACATGATGATGAGAAATNNCA 62
QY      93  slyslsValHlsGlnLeuAlaMetThrValValSerPheHlsGlnValAspTyrThrP 113
Db      63  GAAGAAAGTTCATCAGCTTGCTATGACCGGTGTCAGATTTCATCAGATGATTAATACCTT 122
QY      113 eAspArgAenValLeuSerArgLeuLeuAngLucysArgGluMetLeuHlsGlnIleI 133
Db      123 TCACCGAATGTGTTATCCAGGCTGTTAAATGATGACAGAGATGCTGCACCAATCAT 182
QY      133 eGlnArgHlsLeuThrAlaLysSerHlsGlyArgValAsnAsnValPheAspHisPhe 153
Db      183 TCAGGCCACCTCAGTCCAGATGACATGACGCGTTAAATGATGATGATGATTTTC 242
QY      153 rAspCysGluPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPe 168
Db      243 AGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 288

```

RESULT 37

```

US-10-764-324-5724
/ Sequence 5724, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mamion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ PRIOR FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5724
/ LENGTH: 288

```

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: unsure

/ LOCATION: (58)

/ OTHER INFORMATION: n=A,T,C or G

/ FEATURE:

/ NAME/KEY: unsure

/ LOCATION: (59)

/ OTHER INFORMATION: n=A,T,C or G

US-10-764-324-5724

Alignment Scores:

Pred. No.: 3,65e-45

Length:

288

```

Score: 484.00 Matches: 94
Percent Similarity: 97.92% Conservative: 0
Best Local Similarity: 97.92% Mismatches: 1
Query Match: 50.31% Indels: 1
DB: 19 Gaps: 0

```

US-10-627-571-2 (1-188) x US-10-764-324-5724 (1-288)

```

QY      74  lIeLeuTyrArgAAsnGlnPheAsnGluLeuAlaLeuMetGluLys-PheY 93
Db      3  ATTCTTATAGAGATATATGATTATATCAAGATGACATGATGATGAGAAATNNCA 62
QY      93  slyslsValHlsGlnLeuAlaMetThrValValSerPheHlsGlnValAspTyrThrP 113
Db      63  GAAGAAAGTTCATCAGCTTGCTATGACCGGTGTCAGATTTCATCAGATGATTAATACCTT 122
QY      113 eAspArgAenValLeuSerArgLeuLeuAngLucysArgGluMetLeuHlsGlnIleI 133
Db      123 TCACCGAATGTGTTATCCAGGCTGTTAAATGATGACAGAGATGCTGCACCAATCAT 182
QY      133 eGlnArgHlsLeuThrAlaLysSerHlsGlyArgValAsnAsnValPheAspHisPhe 153
Db      183 TCAGGCCACCTCAGTCCAGATGACATGACGCGTTAAATGATGATGATGATTTTC 242
QY      153 rAspCysGluPheLeuAlaAlaLeuTyrAsnProPheGlyAsnPe 168
Db      243 AGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 288

```

RESULT 38

```

US-09-918-995-35521
/ Sequence 35521, Application US/09918995
/ Publication No. US20030073623A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
/ FILE REFERENCE: 20411-756
/ CURRENT APPLICATION NUMBER: US/09/918,995
/ PRIOR FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: US/09/235,076
/ PRIOR FILING DATE: 1999-01-20
/ NUMBER OF SEQ ID NOS: 38054
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 35521
/ LENGTH: 422
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-918-995-35521

```

Alignment Scores:

```

Pred. No.: 7,34e-27 Length: 422
Score: 324.00 Matches: 62
Percent Similarity: 78.95% Conservative: 28
Best Local Similarity: 54.39% Mismatches: 24
Query Match: 33.68% Indels: 0
DB: 10 Gaps: 0

```

US-10-627-571-2 (1-188) x US-09-918-995-35521 (1-422)

```

QY      6  PheAsnSerLyAsnLeuAlaValGlnAlaGlnLysIleLeuGlyLysMetValSer 25
Db      77  TTCAGCTCAAGAGAGCTGGCACTGCACAGAGAGAAAGCTACTAGTAAGTGGCGGCT 136
QY      26  LysSerIleAlaThrThrLeuIleAspThrSerSerGluValLeuAspGluLeuTyr 45
Db      137  CGCTCTGTGGCTCATCTCTTCATATGATGACACACAGAGAGTGTCTAGTAGCTTAC 196
QY      46  ArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLysIleLeuAsnLeu 65
Db      197  CGTGTCCAGAGAGTACCGCACAGCGGCGCCAGCGGCTGATCAAGACCTG 256
QY      66  lIeLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnAspGlu 85

```



```
Percent Similarity: 96.55% Conservative: 0
Best Local Similarity: 96.55% Mismatches: 1
Query Match: 16.11% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-191-803-330 (1-493)

QY 160 AAlaLeuTyraAnpProPhgIyaenPhelysProHisLeuGlnIyLeuCySaepGlylle 179
DB 491 GCCTTGATATATCCCTTGGAAAATTAAACCTCCTACAGAACTTTGTGACGGCATT 432

QY 180 AenlyMeLeuAspGluGluAsnIle 188
DB 431 AACAAATGTGGATGAAGAACATCA 405

RESULT 42
US-10-029-386-11708
; Sequence 11708, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, David G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11708
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012678.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.51
; OTHER INFORMATION: EST HUMAN HIT: BG10777.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: AB002705.1, EVALUE 5.20e-01
; OTHER INFORMATION: SWISSPROT HIT: P45815, EVALUE 3.90e+00
US-10-029-386-11708

Alignment Scores:
Pred. No.: 7.76e-07 Length: 543
Score: 148.00 Matches: 24
Percent Similarity: 88.64% Conservative: 15
Best Local Similarity: 54.55% Mismatches: 5
Query Match: 15.38% Indels: 0
DB: 16 Gaps: 0

US-10-627-571-2 (1-188) x US-10-029-386-11708 (1-543)

QY 143 GLyArYValAsnAnValPheAspHisPheSerAspCySgluPheLeuAlaLeuTyR 162
DB 3 GGGCGCATCAACACGCTTTAACCACTTGCAGATGTGAGTTCCTCTCCACCTCTAT 62

QY 163 AspProPhgIyaenPhelysProHisLeuGlnIyLeuCySaepGlylleAsnlysmet 182
DB 63 AGCTGTGATGAAGAACTGTAGGCCCAACCTCAAGAGATTGTGAAGGAATCAATTAAGTTG 122

QY 183 LeuAspGluGlu 186
DB 123 CTGATGAAGAAA 134

RESULT 43
US-09-796-692-4917/C
; Sequence 4917, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
```

```
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (77)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-4917

Alignment Scores:
Pred. No.: 0.0192 Length: 538
Score: 109.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.33% Indels: 0
DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x US-09-796-692-4917 (1-538)

QY 169 LysProHisLeuGlnIyLeuCySaepGlylleAsnlysmetLeuAspGluGluAsnIle 188
DB 538 AAACCCCACTTCAAAAATATGTATGTATGATCAACAAATGTGTGATGAAGAACATCA 479

RESULT 44
US-10-040-862-4917/C
; Sequence 4917, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
```

PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4917
LENGTH: 538
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (77)
OTHER INFORMATION: n=A,T,C or G
US-10-040-862-4917

Alignment Scores:
Pred. No.: 0.0192 Length: 538
Score: 109.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.33% Indels: 0
DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-040-862-4917 (1-538)

QY 169 LysEProHisLeuGlnIleuCyAspGlyIleAsnLysMetLeuAspGluGluAsnIle 188
Db 538 AAACCCACTTACAAAATGATGTGATGATCAACAAAATGTTGATGAAGAACATCA 479

RESULT 45

US-10-057-475B-4917/C
Sequence 4917, Application US/10057475B
Publication No. US2004002068A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordonez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4917
LENGTH: 538
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(538)
OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-4917

Alignment Scores:
Pred. No.: 0.0192 Length: 538
Score: 109.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.33% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-057-475B-4917 (1-538)

QY 169 LysEProHisLeuGlnIleuCyAspGlyIleAsnLysMetLeuAspGluGluAsnIle 188
Db 538 AAACCCACTTACAAAATGATGTGATGATCAACAAAATGTTGATGAAGAACATCA 479

Search completed: July 28, 2005, 19:36:11
Job time: 671 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 19:13:42 ; Search time 2996 Seconds

(without alignments)
2388.546 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962

Sequence: 1 MATDVNSKRLAVQAOAKLIL.....KPHLOKLCDCGIMMLDEBNI 188

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 6847988

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=x1h
-O=/cgm2_1/USPTO.spool_h/US10627571/runat_27072005_154720_27509/app_query.fasta_1.327
-DB=EST -OPMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCAL=GN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10627571.@CGN 1.1 3437 @runat_27072005_154720_27509 -NCPU=6 -ICPU=3
-NO MMP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_g861: *
9: gb_g862: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	954	99.2	737	7	CN305329 170006000
2	954	99.2	780	6	CB990165
3	954	99.2	806	7	CF593784
4	954	99.2	813	4	BI181214
5	954	99.2	920	7	CF272384
6	950	98.8	712	7	CN305327
7	950	98.8	731	6	CD521721
8	950	98.8	1055	5	BM919223
9	949	98.6	718	1	AL549492

10	949	98.6	885	1	AL550457
11	949	98.6	1610	3	AL527566
12	949	98.6	1733	3	CR611795
13	949	98.6	1733	3	CR615621
14	949	98.6	1754	5	CR615131
15	947	98.4	896	5	BQ424670
16	947	98.4	998	7	CN648259
17	945	98.2	1045	1	AL554168
18	934	97.1	779	7	CR548728
19	933	97.0	876	4	BI752550
20	927	96.4	773	5	BP433798
21	926	96.4	669	3	CN789950
22	924	96.0	942	6	CA984413
23	923	95.9	893	4	BI759027
24	920	95.6	832	2	BF607236
25	920	95.6	841	6	CA469176
26	920	95.6	865	4	BI525793
27	918	95.4	559	1	AL558873
28	918	95.4	580	5	BP302623
29	916	95.2	651	4	BG146795
30	916	95.2	816	4	AU135377
31	914	95.0	1542	3	AK090316
32	912	94.8	908	4	BI688217
33	909	94.5	766	4	BI683693
34	907	94.3	772	4	BI103517
35	907	94.3	785	6	CB991862
36	905	94.1	1836	3	AK087478
37	904	94.0	622	6	CD701806
38	902	93.8	656	4	BI687787
39	901	93.7	707	7	CR557522
40	900	93.6	566	7	CV023086
41	898	93.3	779	1	AL542926
42	893	92.8	747	6	CD520304
43	892	92.7	671	6	BY743355
44	891	92.6	739	2	BF140519
45	889	92.4	692	7	CR771358
46	885	92.0	711	4	BG972935
47	885	92.0	834	5	BK405197
48	884	91.9	580	3	BP228279
49	884	91.9	580	3	AK090350
50	879.5	91.4	757	6	CD520319
51	879	91.4	697	1	AV713586
52	874	90.9	596	6	CA843219
53	874	90.9	673	5	BP459271
54	872.5	90.7	593	7	CV027255
55	868	90.2	672	7	CN305328
56	865	89.9	762	5	BU124386
57	863	89.7	588	6	CD723203
58	863	89.7	837	4	BI690706
59	856	89.0	837	4	BG534704
60	852	88.6	548	1	AA116718
61	850	88.4	648	4	BG146696
62	848	88.1	863	4	BI686259
63	844	87.7	572	7	CK903031
64	844	87.7	600	6	CA526496
65	842.5	87.6	942	7	CK805254
66	840.5	87.4	676	1	AL642789
67	840.5	87.4	835	5	BK718081
68	840.5	87.4	864	5	BK733894
69	840.5	87.4	900	5	BK743252
70	840	87.3	784	4	CG640021
71	839.5	87.3	849	5	BK849289
72	838	86.9	605	7	CR771295
73	836	86.9	616	6	CB545575
74	835	86.8	734	4	CF737286
75	825.5	85.7	788	7	CR444882
76	824.5	85.7	788	7	CR444882
77	818.5	85.1	688	6	BY743855
78	816	84.8	744	4	BI760235
79	814	84.6	671	6	CB551156
80	812.5	84.5	640	6	CD639093
81	809	84.1	607	6	CB552439
82	808	84.0	863	6	CD522233

AL550457	AL550457
AL527566	AL527566
CR611795	full-length
CR615621	full-length
CR615131	full-length
BQ424670	AGENCOURT
CN648259	ILLUMINAGEN
AL554168	AL554168
CR548728	DRFZP469E
BI752550	603021969
BP433798	BP433798
CN789950	4124483 B
CA984413	AGENCOURT
BI759027	603042814
BF607236	MT1_00017
CA469176	AGENCOURT
BI525793	602924756
AL558873	AL558873
BP302623	BP302623
BG146795	mdb95H09.
AU135377	AU135377
AK090316	Mus muscu
BI688217	603314718
BI683693	603306241
BI103517	602889058
CB991862	AGENCOURT
AK087478	Mus muscu
CD701806	EST18330
BI687787	603314107
CR557522	DKEZP469E
CV023086	5463 Full
AL542926	AL542926
CD520304	AGENCOURT
BY743355	BY743355
BF140519	601787536
CR771358	DKEZP469E
BG972935	602840605
BK405197	BK405197
BP228279	BP228279
AK090350	Mus muscu
CD520319	AGENCOURT
AV713586	AV713586
CA843219	i55C03.Y
BP459271	BP459271
CV027255	5464 Full
CN305328	170005999
BU124386	603148153
CD723203	6119607.Y
BI690706	603314314
BG534704	602553867
AA116718	mg24002.r
BG146696	mdb94H09.
BI686259	603312767
CK903031	i55C03.X
CA526496	8012-74 M
CK805254	AGENCOURT
AL642789	AL642789
BK718081	BK718081
BK733894	BK733894
BK743252	BK743252
CG640021	OST372269
BK849289	BK849289
CR771295	DKEZP469M
CB545575	NMSPO071
CF737286	U-M-HDO-
BG444882	602506586
BY743855	BY743855
BI760235	603045171
CB551156	NMSPO066
CD639093	AGENCOURT
CB552439	NMSPO048
CD522233	AGENCOURT

83	807	83.9	622	6	CB553311	MMSP0007
84	801	83.3	588	7	CR787140	DKEZD4695
85	799.5	83.1	771	6	CD521069	AGENCOURT
86	799.5	83.1	829	7	CK804285	AGENCOURT
87	799	83.1	674	2	BB667314	BB667314
88	795	82.6	821	1	AJ451661	AJ451661
89	795	82.6	975	2	BF781663	602104381
90	794	82.5	652	6	CD638280	AGENCOURT
91	793	82.4	583	5	BP345938	BP345938
92	793	82.4	648	7	CN053290	Salamaend
93	789	82.0	681	6	CD683714	EST234 hu
94	782	81.3	909	2	BE913315	60166522
95	780	81.1	661	7	CF737920	UI-M-HD-
96	778	80.9	766	1	AJ453373	AJ453373
97	777.5	80.8	648	1	AL888174	AL888174
98	772	80.2	582	5	BP261675	BP261675
99	770.5	80.1	680	4	BJ068640	BJ068640
100	757	78.7	517	6	CD372026	UI-R-G00-

ALIGNMENTS

RESULT 1
CN305329 737 bp mRNA linear EST 16-MAY-2004
LOCUS 17000600027114 GRN_PPREHP Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN305329
ACCESSION CN305329
VERSION CN305329.1 GI:47321743
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 737)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Mursge, J., Fisk, G. J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M. S., Mandalam, R.,
Lebkowski, J. and Stanton, L. W.

TITLE
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)

JOURNAL
COMMENT
Contact: Brandenberger R
Regenerative Medicine
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@regeron.com

FEATURES
source
1..737
Location/Qualifiers
Insert Length: 737 Std Error: 0.00.

ORIGIN
Pred. No.: 6 73e-105 Length: 737
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: 7 Gaps: 0
US-10-627-571-2 (1-188) x CN305329 (1-737)
1 MetalThAspValPheAsnSerLysAsnLeuAlValGlnAlaGlnLysLysIleLeu 20
|||||

Db	107	ATGGCCACAGATGCTTTAATTCAGAAAAAAGCTGGCGTTCCAGGACAAAGAAAGATCTTG	166
Qy	21	GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal	40
Db	167	GGTAAATGCTCCAAATCCATCCGACACCACTTAATAGACAAAGTGTAGGGT	226
Qy	41	LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLysLysGluIleLys	60
Db	227	CTGGATGACCTTACAGAGTCCAGGAGTACACCCAAACAAAGAGGACAGAAAG	286
Qy	61	LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnGln	80
Db	287	ATCATCAAGAACCTCATCAAGACAGTGCATCAAGCTGGCCATCTTTATFGAATATCG	346
Qy	81	PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla	100
Db	347	TTTAAATCAAGATGACCTGACATGTAGTGAATTAAGAAAGATTATCATGCTTGT	406
Qy	101	MetThrValIleSerPheIleGlnValAspTyThrPheAspArgAsnValLeuSerArg	120
Db	407	ATGACCGTGGTCAAGTTTCCATCAGGTGATTTACCTTTGACCGGAATGTGTTATCCAG	466
Qy	121	LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys	140
Db	467	CTGTTAAATGAATGCAGAGATGCTGCACCAATCATTCACGCGCACCTCAGTCCAA	526
Qy	141	SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle	160
Db	527	TCACATGACGGGTAAATGATGTGTTGATCTTTTCAGATTTGAAATTTTGGCTGCG	586
Qy	161	LeuTyAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn	180
Db	587	TTGTATATCTCTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATGATCATC	646
Qy	181	LysMetLeuAspGluGluAsnIle	188
Db	647	AAATGTTGATGAAGAAACATATA	670

RESULT 2
CB990165 780 bp mRNA linear EST 01-MAY-2003
LOCUS AGENCOURT 13902271 NIH MGC 147 Homo sapiens cDNA clone
DEFINITION IMAGE:30347855 5', mRNA sequence.
ACCESSION CB990165
VERSION CB990165.1 GI:30284685
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 780)
NIH-MGC http://mgs.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDAM389 row: 5 column: 24
High quality sequence stop: 635.
FEATURES
source
1..780
Location/Qualifiers
Insert Length: 780
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: 7 Gaps: 0
US-10-627-571-2 (1-188) x CN305329 (1-737)
1 MetalThAspValPheAsnSerLysAsnLeuAlValGlnAlaGlnLysLysIleLeu 20
|||||


```

Db      438 ATGACCGGTGTCAGTTTCCATCATCAGTGGATTATACCTTGGACCGGAGATGTTATCCAGG 497
Qy      121 LeuLeuAsnGluCyArArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIys 140
Db      498 CTGTTAAATGATATCGAGAGATGCTGCACCAATATTCAGGCCACCTCAGCTGCAGG 557
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
Db      558 TCACATGAGACGGGTATATATATGCTTTCATCATTTTTCAGATTGTGATTTTGGCTGCC 617
Qy      161 LeuTyrAsnProPheGlyAsnPhelyProHisLeuGlnLysLeuCysAspGlyIleAsn 180
Db      618 TTGTATAATATCCTTTGGGATTTTAAACCCACTTACAAACTATGATGTATCAAC 677
Qy      181 LysMetLeuAspGluGluAsnIle 188
Db      678 AAAATGTTGGATGAGAGAACATA 701

RESULT 4
BI819214      813 bp      mRNA      linear      EST 04-OCT-2001
LOCUS      603034630F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175702 5',
DEFINITION      mRNA sequence.
ACCESSION      BI819214
VERSION      BI819214.1 GI:15930764
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1437 row: 1 column: 07
High quality sequence stop: 809.

FEATURES
source      Location/Qualifiers
1..813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5175702"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 Kb,
insert size range 1-3 Kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      7.68e-105      Length:      813
Score:      954.00      Matches:      187
Percent Similarity:      99.47%      Conservative:      0
Best Local Similarity:      99.47%      Mismatches:      1
Query Match:      99.17%      Indels:      0
Db:      4      Gaps:      0

```

```

US-10-627-571-2 (1-188) x BI819214 (1-813)
Qy      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeu 20
Db      98 ATGGCCACAGAGATCTTTAATTCAAAACCTGGCCGTTCAAGCAAAAAGAAATCTTG 157
Qy      21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db      158 GGTAAATATGATCTCCAAATCCATCCGCCACCTTAATGACGACAAAGTAGTAGGG 217
Qy      41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGlnAlaGluLys 60
Db      218 CTGATAGAGCTTACAGAGTGACAGGAGTGACACCAAAACAGAGAGGACAGAAAG 277
Qy      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgSerAsnGln 80
Db      278 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAG 337
Qy      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db      338 TTTAATCAAGATGAGCTGACATTGATGAGAAATTTAAGAAAGAAAGTTATCAGCTTGT 397
Qy      101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
Db      398 ATGACCGGTGTCAGTTTCCATCAGGTGATTTATACCTTTGACCGGAATGTATTCCAGG 457
Qy      121 LeuLeuAsnGluCyArArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIys 140
Db      458 CTGTTAAATGAATGAGAGAGATGCTGCACCAATCATTCAGCGCCACTCCTCCAGG 517
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
Db      518 TCACATGAGACGGGTATATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 577
Qy      161 LeuTyrAsnProPheGlyAsnPhelyProHisLeuGlnLysLeuCysAspGlyIleAsn 180
Db      578 TTGTATAATCCTTTGGGATTTTAAACCCACTTACAAACTATGATGTATCAAC 637
Qy      181 LysMetLeuAspGluGluAsnIle 188
Db      638 AAAATGTTGGATGAGAGAACATA 661

RESULT 5
CF272384      920 bp      mRNA      linear      EST 13-AUG-2003
LOCUS      AGENCOURT.15178772 NIH_MGC_192 Homo sapiens cDNA clone
DEFINITION      IMAGE:30512832 5', mRNA sequence.
ACCESSION      CF272384
VERSION      CF272384.1 GI:33628296
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genetics
National Cancer Institute / NIH
Bldg. 31 Rm10K07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Agencourt
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDAM573 row: m column: 01
High quality sequence stop: 598.

FEATURES
Location/Qualifiers

```


Source

1. .920
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAG:30512832"
 /tissue_type="Brain glioblastoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lhb="NH1 MGC_192"
 /note="Vector: pExpress1; Site 1: SmaI; Site 2: NotI; The library was constructed by reverse transcription of 1 ug mRNA using the oligo dt primer GCGGCGGCC(7)20 and an RNaseH⁺ MMLV reverse transcriptase. Second strand synthesis was carried out by standard methods. The cDNA was size selected by agarose gel for > 1.2 kb, digested with Not I and directionally cloned into the vector Express-1 at the SmaI/NotI sites. DNA from the primary library was used for in vitro transcription from the T7 promoter to produce biotinylated RNA transcripts. These biotinylated transcripts, along with blocking oligos to the poly-A, multiple cloning site and primer regions, were hybridized with single stranded circles produced by phageid production from the primary library to a Cot value of 10-20. Streptavidin/phenol extraction was utilized to remove DNA:RNA hybrids leaving un-hybridized single stranded circles which were repaired by primer extension and transformed back into E. coli resulting in the normalized library. Average insert size 2.0 kb. 3' linker/adaptor sequence GCGGCGGCC(7)20. This library was constructed by Agencourt Bioscience."

ORIGIN

Alignment Scores:
 Pred. No.: 9.05e-105 Length: 920
 Score: 954.00 Matches: 187
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 99.17% Indels: 0
 DB: 7 Gaps: 0

US-10-627-571-2 (1-188) x CF272384 (1-920)

QY 1 MetAlaTharApVaAlPheAsnSerLyAsnLeuAlaVaIGlnAglNylsYlLeu 20
 Db 98 ATGGCCACAGATGCTTTAATTCGAAACCTGCGCTTACGACCAAGAAAGATCTTG 157
 QY 21 GilylYMeValSerlySerIleAlaThrLeuIleAspAPThrSerSergIuVal 40
 Db 158 GGTAAATGGTGTCAATCATCGCCACCTTATAGACGACAAAGTAGAGGTG 217
 QY 41 LeuAspGluLeuTyArgValThraGluTyThrGlnAsnYlsgIuAgluYs 60
 Db 218 CTGGATAGCTCTACAGAGTACGACGAGTAGACCCAAACAAAGAGGACGAGAG 277
 QY 61 LysIlelyAsnleuIlelySerThrValIlelyleuAlaIleleuTyArgAsnAenGln 80
 Db 278 ATCATCAAGACCTTCACCAAGACGATCATCAAGCTGCGCTTCTTATAGAAATTAACAG 337
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluYsPheYslylYsValHlsgInLeuAla 100
 Db 338 TTTTATCAAGATAGCTACGATTCGATGAGAAATTTTAAAGAAAGTTTCATCAGCTTGTCT 397
 QY 101 MetThrValAlaSerPheHlsgInValAspTyThrPheAspArgAsnValLeuSerArg 120
 Db 398 ATGACCGTGTCACTTCATCAGTGTGATTAATCCTTACCGGAATGTGTTATCCAGG 457
 QY 121 LeuLeuAsnGluCyAsArgGluMetLeuHlsgInIleleGlnArgHlsgleuThraAlYs 140
 Db 458 CTGTTAATATGATGACAGAGATGCTCACCAATCTTCAGGCGCCACCTCAGCTCCAG 517
 QY 141 SerHlsgIyArgValaAsnAsnValPheAspHlsgPheSerAspCySgluPheLeuAla 160
 Db 518 TCACATGAGAGGGTATTAATGCTTTCATTCATTTTCAATTTGTGAATTTTGGCTGCC 577

QY 161 LeuTyAsnProPheGlyAsnPhelysProHlsgLeuGlnLysLeuCyAspGlyIleAsn 180
 Db 578 TTGATATATCTCTTTGGGAAATTTTAAACCCACTTACAAAAACTATGATGATTCAC 637
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 638 AAATGTGTGATCAAGAGAAACATA 661

RESULT 6
 CN305327 712 bp mRNA linear EST 16-MAY-2004
 LOCUS 1700053262591 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
 DEFINITION CN305327
 ACCESSION CN305327
 VERSION CN305327.1 GI:47321741
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 712)
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muraige, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.
 Transcriptionome characterization elucidates signaling networks that control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 CONTACT: Brandenberger R
 Regenerative Medicine
 Genon Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@genon.com
 Insert Length: 712 Std Error: 0.00.
 Location/Qualifiers

FEATURES

1. .712
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, cell lines H1, H7, and H9"

Alignment Scores:
 Pred. No.: 1.96e-104 Length: 712
 Score: 950.00 Matches: 186
 Percent Similarity: 99.47% Conservative: 1
 Best Local Similarity: 98.94% Mismatches: 1
 Query Match: 98.75% Indels: 0
 DB: 7 Gaps: 0

US-10-627-571-2 (1-188) x CN305327 (1-712)

QY 1 MetAlaTharApVaAlPheAsnSerLyAsnLeuAlaVaIGlnAglNylsYlLeu 20
 Db 81 GTGGCCACAGATGCTTTAATTCGAAACCTGCGCTTACGACCAAGAAAGATCTTG 140
 QY 21 GilylYMeValSerlySerIleAlaThrLeuIleAspAPThrSerSergIuVal 40
 Db 141 GGTAAATGGTGTCAATCATCGCCACCTTATAGACGACAAAGTAGAGGTG 200
 QY 41 LeuAspGluLeuTyArgValThraGluTyThrGlnAsnYlsgIuAgluYs 60
 Db 201 CTGGATAGCTCTACAGAGTACGACGAGTAGACCCAAACAAAGAGGACGAGAG 260
 QY 61 LysIlelyAsnleuIlelySerThrValIlelyleuAlaIleleuTyArgAsnAenGln 80
 Db 261 ATCATCAAGACCTTCATCAAGACGATCATCAAGCTGCGCTTCTTATAGAAATTAACAG 320

```

QY      81 PheASglnAAspGluLeuValaleuMetGluLysPheLysLysValHisGlnLeuA1a 100
      |||
Db      321 TTTAAATCAAGATAGCTAGCATTTGATGAGAAATTTTAAAGAAAGTTCTATCAGCTTGC 380
QY      101 MetThrValAlaSerPheHisGlnValAspTyrThrPheAspArgAnValIleuSerArg 120
      |||
Db      381 ATGACCGGTGTCAGTTCCATCATCGATGATTATACCTTTGACCGGAATGTTATCCAG 440
QY      121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
Db      441 CTGTAAATGAATGACGAGAGATGCTGCACCAATCATTTAGCGCCACCTCATGCGCAAG 500
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla1a 160
      |||
Db      501 TCACATGACGCGGTTAATATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 560
QY      161 LeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
      |||
Db      561 TTGTATTAATCTCTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGATCAAC 620
QY      181 LysMetLeuAspGluGluAsnIle 188
      |||
Db      621 AAAATGTTGATGAAGAACATTA 644

RESULT 7
LOCUS   CD521721 731 bp mRNA linear EST 06-JUN-2003
DEFINITION AGENCOURT_14353034 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:30413002 5', mRNA sequence.
ACCESSION CD521721
VERSION   CD521721.1 GI:31453439
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   Contact: Daniela S. Gerhard, Ph.D.
COMMENT   Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: Narayan Bhat
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnl.gov
          Plate: NDCM207 row: e column: 11
          High quality sequence stop: 603.
          Location/Qualifiers
            1..731
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:30413002"
              /issue_type="Pooled"
              /lab_host="DH10B (Ti phage-resistant)"
              /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcagc);
              Site 2: SfiI (ggcgctcgcc); Library is oligo-dT primed
              and directionally cloned. PBMC - Peripheral Blood
              Mononuclear Cells. RNA was pooled from 3/6hour stimulation
              with PMA adn Ionomycin. 5' and 3' adaptors were used in
              cloning as follows: 5' adaptor sequence:
              5'-CAGCGCATTTATGCG-3' and 3' adaptor sequence:
              5'-ATTCTAGAGCGCGAGCGCGCATG-dt(30)BN-3' (where B = A,
              C, or G and N = A, C, G, or T). Average insert size 1.69
              kb (range 0.70-5.0 kb). 15/15 colonies contained inserts

```

ORIGIN

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC library."

Alignment Scores:

Pred. No.:	2,03e-104	Length:	731
Score:	950.00	Matches:	186
Percent Similarity:	99.47%	Conservative:	1
Best Local Similarity:	98.94%	Mismatches:	1
Query Match:	98.75%	Indels:	0
DB:	6	Gaps:	0

US-10-627-571-2 (1-188) x CD521721 (1-731)

```

QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeu 20
      |||
Db      30 GTGGCCACAGATGTCCTTTAATTCCAAAAACCTGGCGCTTCAGGCACAAAAGAAATCTTG 89
QY      21 GlyMetValSerLysSerIleAlaThrLeuIleAspAspThrSerGluVal 40
      |||
Db      90 GGTAAATGTGTCCAAATTCATCCGACACCTTAATGACGACAAAGATGTAGG 149
QY      41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLys 60
      |||
Db      150 CTGGAATGAGCTCTACAGAGTGACCGAGGAGTACACCCAAAACAAGAGGAGGACAGAAAG 209
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
      |||
Db      210 ATCATCAAGAACCTCATCAAGACAGTCAACAGCTCGGCATCTTATGGAATATACAG 269
QY      81 PheAsnGlnAAspGluLeuValaleuMetGluLysPheLysLysValHisGlnLeuA1a 100
      |||
Db      270 TTTAAATCAAGATAGCTAGCATTTGATGAGAAATTTTAAAGAAAGTTCTATCAGCTTGC 329
QY      101 MetThrValAlaSerPheHisGlnValAspTyrThrPheAspArgAnValIleuSerArg 120
      |||
Db      330 ATGACCGGTGTCAGTTCCATCATCGATGATTATACCTTTGACCGGAATGTTATCCAG 389
QY      121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
Db      390 CTGTAAATGAATGACGAGAGATGCTGCACCAATCATTCACCGCACCTCATGCAAG 449
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla1a 160
      |||
Db      450 TCACATGACGCGGTTAATATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 509
QY      161 LeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
      |||
Db      510 TTGTATTAATCTCTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGATCAAC 569
QY      181 LysMetLeuAspGluGluAsnIle 188
      |||
Db      570 AAAATGTTGATGAAGAACATTA 593

RESULT 8
LOCUS   BM919223 1055 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6715690 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748442
IMAGE:5748442 5', mRNA sequence.
ACCESSION BM919223
VERSION   BM919223.1 GI:19369602
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   Contact: Robert Strausberg, Ph.D.
COMMENT   Email: cgabbs-r@mail.nih.gov

```


QY 82 AaenGlnaSpGluLeuAlaLeuMetGluYrPhelysLysValHisGlnLeuAlaMet 101
 DB 243 AATCAAGATGAGCTAGCATTTGATGAGAAATTTAAAGAGAAATTCATCGCTTGCCTATG 302
 QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
 DB 303 ACCGTGTCAGTTTCCATCGATGAGTGAATTAACCTTTGACCGGAATGTGATTCAGAGCTG 362
 QY 122 LeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
 DB 363 TTAATGAATGACGAGAGATGCTGCACCAATCATTCAGCGCCACCTCACTGCGCAAGTCA 422
 QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeu 161
 DB 423 CATTGACGGGTATATATATGTTTGTATCATTTTTCAGATTTGTAATTTTGGCTGCTTGG 482
 QY 162 TyrAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleAsnLys 181
 DB 483 TATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTATGTATCAACAA 542
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 543 ATGTTGATGAAGAGAACATA 563
 RESULT 10
 AL550457 885 bp mRNA linear EST 25-MAR-2004
 LOCUS AL550457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION
 AL550457
 VERSION AL550457
 KEYWORDS GI:45750828
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31272274.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by life technologies, a
 division of invitrogen. This sequence belongs to sequence cluster
 3485.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CSODI057AA04QPl&c=3485.f.
 FEATURES
 source
 1..885
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI057YA07"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,47e-104 Length: 885
 Score: 949.00 Matches: 186
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 1

Query Match: 98.65% Indels: 0
 DB: 1 Gaps: 0
 US-10-627-571-2 (1-188) x AL550457 (1-885)
 QY 2 AaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGly 21
 DB 3 GCCACGATGCTTTTAAATTCAAAACCTGGCGCTTACGGCACAAAGAGATCTTGGGT 62
 QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
 DB 63 AAAATGGTGTCCAAATTCATCCACCACTTAAATAGACACAAAGATGATGAGTGTCTG 122
 QY 42 AaPGLuLeuTyrArgValThrArgLysThrGlnAsnLysValGlnAlaGlnLysLys 61
 DB 123 GATGACCTTACAGAGTGAACGAGAGTACACCCAAACAGAGAGGACGAGAAATGC 182
 QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPhe 81
 DB 183 ATCAAGAACTCATCAAGACAGTATCAAGTGGCCATTTCTTATAGGATATACGTTT 242
 QY 82 AaenGlnaSpGluLeuAlaLeuMetGluYrPhelysLysValHisGlnLeuAlaMet 101
 DB 243 AATCAAGATGAGCTAGCATTTGATGAGAAATTTAAAGAGAAATTCATCGCTTGCCTATG 302
 QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
 DB 303 ACCGTGTCAGTTTCCATCGATGAGTGAATTAACCTTTGACCGGAATGTGATTCAGAGCTG 362
 QY 122 LeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
 DB 363 TTAATGAATGACGAGAGATGCTGCACCAATCATTCAGCGCCACCTCACTGCGCAAGTCA 422
 QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeu 161
 DB 423 CATTGACGGGTATATATATGTTTGTATCATTTTTCAGATTTGTAATTTTGGCTGCTTGG 482
 QY 162 TyrAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleAsnLys 181
 DB 483 TATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTATGTATCAACAA 542
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 543 ATGTTGATGAAGAGAACATA 563
 RESULT 11
 AL527566 897 bp mRNA linear EST 24-MAR-2004
 LOCUS AL527566 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION
 AL527566
 VERSION AL527566
 KEYWORDS GI:45702664
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 13, 2001 this sequence version replaced gi:31065417.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized. Library
 was constructed by life technologies, a
 division of invitrogen. This sequence belongs to sequence cluster
 3485.f
 For more information about this cluster, see

FEATURES <http://www.genoscope.cns.fr/cdna?S=CS0DC024AA120P1&c=3485.f>.
Location/Qualifiers

source
1. .897
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC024YA23"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
Pred. No.: 3,53e-104 Length: 897
Score: 949.00 Matches: 186
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 98.65% Indels: 0
DB: 1 Gaps: 0

US-10-627-571-2 (1-188) x AL527566 (1-897)

QY 2 AATThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLysLysIleLeuGly 21
DB 3 GCCACAGATGCTTTTAATTCCAAAAACCTGGCCGTTCCAGCACAAGAAAGATCTGGGT 62
QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValIleu 41
DB 63 AAAATGGTGCCTCAATCCATCGCCACCACTTAATAGACACCAAGTACTGAGTGGCTG 122
QY 42 AspGluLeuTyArgValThrArgGluTyThrGlnAsnLysGlnAlaGlnLysLys 61
DB 123 GATGAGCTCTACAGAGTGACAGAGAGTACACCCAAACAGAGAGGAGGAGAAATC 182
QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnGlnPhe 81
DB 183 ATCAAGACCTCATCAAGACGATCATCAAGCTGGCCATCTTATAGAAATATCAGTTT 242
QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
DB 243 AATCAAGATGAGTATGATGAGAAATTTAAGAGAAAGTATCATCGCTTGTATG 302
QY 102 ThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeu 121
DB 303 ACCGTGCTAGTTTCATCATCGTGGATTAACCTTTGACCGGAATGTGTTATCCAGGCTG 362
QY 122 LeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
DB 363 TTTAAATGAATGCAGAGAGATGCTGCACCAATATCATTCAGCGCCACTGCAAGTCA 422
QY 142 HisGlnArgValAsnAsnValPheAspHisPheSerAspCyArgLysPheLeuAlaLeu 161
DB 423 CATGACGGGTTAATATGTTTGAATCAATTTTTCAGATTGTGAATTTTGGCTGCTTG 482
QY 162 TyAsnProPheGlnAsnPheLysProHisLeuGlnLysLeuCyAspGlyIleAsnLys 181
DB 483 TATTAATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATGATATCAACAA 542
QY 182 MetLeuAspGluGlnLysAsnIle 188
DB 543 ATGTTGATGAGAGAGAACATTA 563

RESULT 12

CR611795 1610 bp mRNA linear HTC 21-JUL-2004
LOCUS Full-length cDNA clone CS0D1057YA07 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR611795
VERSION CR611795.1 GI:50492602
KEYWORDS HTC, CNSLT_cDNA.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 1610)
Li, W.B., Gruber, C., Jesse, J., and Polyes, D.
Full-length cDNA libraries and normalization

Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue

REMARK

REFERENCE

Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

Location/Qualifiers
1. .1610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1057YA07"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:
Pred. No.: 7.7e-104 Length: 1610
Score: 949.00 Matches: 186
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 98.65% Indels: 0
DB: 3 Gaps: 0

US-10-627-571-2 (1-188) x CR611795 (1-1610)

QY 2 AATThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLysLysIleLeuGly 21
DB 3 GCCACAGATGCTTTTAATTCCAAAAACCTGGCCGTTCCAGCACAAGAAAGATCTGGGT 62
QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValIleu 41
DB 63 AAAATGGTGCCTCAATCCATCGCCACCACTTAATAGACACCAAGTACTGAGTGGCTG 122
QY 42 AspGluLeuTyArgValThrArgGluTyThrGlnAsnLysGlnAlaGlnLysLys 61
DB 123 GATGAGCTCTACAGAGTGACAGAGAGTACACCCAAACAGAGAGGAGGAGAAATC 182
QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnGlnPhe 81
DB 183 ATCAAGACCTCATCAAGACGATCATCAAGCTGGCCATCTTATAGAAATATCAGTTT 242
QY 142 HisGlnArgValAsnAsnValPheAspHisPheSerAspCyArgLysPheLeuAlaLeu 161
DB 423 CATGACGGGTTAATATGTTTGAATCAATTTTTCAGATTGTGAATTTTGGCTGCTTG 482

QY	162	TyrAsnProPheGlyAsnPhenylProHisIleuGlnYsleuCyAsnGlyTlleAsnYls	181
Db	483	TATATACCTTTGGGAATTTTAAACCCCACTTACAAAATATATGTGATGATCAACAA	542
QY	182	MetLeuAepGluGluAsnIle	188
Db	543	ATGTTGATGAAGACACATA	563
RESULT 13			
LOCUS	CR615621	1733 bp	mRNA
DEFINITION	full-length cDNA clone CS0C024YA23 of Neuroblastoma Cot		
ACCESSION	CR615621		
VERSION	CR615621.1	GI:50496428	
KEYWORDS	HTC; CNS1T_cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1733)		
TITLE	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.		
JOURNAL	full-length cDNA libraries and normalization		
REMARK	Unpublished		
	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradey Avenue		
REFERENCE	2 (bases 1 to 1733)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
COMMENT	1st strand cDNA was primed with a NotI-cligo(dt) primer. Five primase end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and BcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.		
FEATURES	Location/Qualifiers		
source	1..1733		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0C024YA23"		
	/tissue_type="Neuroblastoma Cot 25-normalized"		
	/plasmid="pCMVSPORT_6"		
ORIGIN			
Alignment Scores:			
Pred. No.:	8.5e-104	Length:	1733
Score:	949.00	Matches:	186
Percent Similarity:	99.47%	Conservative:	0
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	98.65%	Indels:	0
Db:	3	Gaps:	0
US-10-627-571-2 (1-188) x CR615621 (1-1733)			
QY	2	AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnYsleuGly	21
Db	3	GCCACAGATGCTTTAATTCACAAAACCTGCGCGTTCCAGGACACAAAGAAATCTTGGGT	62
QY	22	LysMetValSerLysSerLleAlaThrThrIleuIleAaPAPThrSerSeGluValIleu	41
Db	63	AAATATGCTGCACAAATTCATGCGCACCACTTATATAGACGACACAGATGAGGTGCTG	122
QY	42	AspGluLeuYrArgValThrArgGlnYrThrGlnAsnLysGluAlaGluLys	61
Db	123	GATGAGCTCTACAGAGTGACCGAGGAGTACACCCAAAACAAAGAGGACGAGAAGATC	182
QY	62	IlleYsAsnLeuIlleYsThrValIlleYsLeuAlaIleLeuYrArgAsnAsnGlnPhe	81

Db	183	ATCAAGAACCTTCATCAGACAGCTCATGACGTGGCCANTCTTTATAGGAATAATACGTTT	242		
QY	82	AaenqlnspgltuleuAlaLeuMeGltulysPheylslylsvAlaHlsglLeuAlaMet	101		
Db	243	AATCAAGATGAGCTAGCATGTATGAGAGAAATTTAAGAAAGATTCACTACCTGCTATG	302		
QY	102	ThrValValIsarpPheHlsglNValAspPyrThrPheAspArganValLeuSerArgIeu	121		
Db	303	ACCGGTGCATCTTCCATCAGCTGATTAATACCTTTGACCGGAATGVTATCCAGGCTG	362		
QY	122	LeuAenqluCyBAArglUmeLeuHlsglNlelleglNArgHlsleuThraAlaIysSer	141		
Db	363	TTAAATGAATCAGAGAGATCTGTGACCAATATCACTACGCCACCTCATCTCCAGACGA	422		
Db	142	HlsglArgValAaenAenValPheAspHlspheSerAspCyGltPheLeuAlaAlaLeu	161		
QY	423	CATGACGGGTAAATTAATGTGTGATCATTTTTCAGATGTGAATTTTGGCTGCCTTG	482		
Db	162	TyrAsnProPheGlyAenPheIysProHlslsleuGlnIylsLeuCyBAAspGlyTlAasnIys	181		
QY	483	TATATACCTTTTGGAAATTTTAAACCCCATTAACAAAACTAATGATGATCAACAA	542		
Db	182	MetLeuAspGltugluAenlle	188		
Db	543	ATGTTGATGAGAGACACTA	563		
RESULT 14					
LOCUS	CR615131	1754 bp	mRNA linear HTC 21-JUL-2004		
DEFINITION	full-length cDNA clone CSDD1081Y011 of Placenta Cot 25-normalized of Homo sapiens (human).				
ACCESSION	CR615131				
VERSION	CR615131.1	GI:50495938			
KEYWORDS	HTC; CNSLT cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1754)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NctI-oligo(dT) primer. Five primers end enriched. double-strand cDNA was digested with Nct I and cloned into the Nct I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..1754				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSDD1081Y011"				
	/tissue type="Placenta Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Alignment Scores:					
Pred. No.:	8.64e-104	Length:	1754		
Score:	949.00	Matches:	186		
Percent Similarity:	99.47%	Conservative:	0		
Best Local Similarity:	99.47%	Mismatches:	1		
Query Match:	98.65%	Indels:	0		
DB:	3	Gaps:	0		

US-10-627-571-2 (1-188) x CR615131 (1-1754)

```

QY 2 AlAThAspVa1PheAsnSerLyAsnLeuAlaValGlnAglNylsYlLeuGly 21
   |||||
DB 3 GCCACAGATGTCCTTAATCCAAAAACCTGGCCGCTCAGGACCAAAAGATCTTG 62
   |||||
QY 22 LyMeTAspSerLySerLleAlaThrThLeuLeuAspPheSerSerGluValLeu 41
   |||||
DB 63 AAAATGGTGTCCAAATCATCGCACCACTTAATACACACAAAGTAGAGAGTCTG 122
   |||||
QY 42 AspGluLeuTyraValThrArgGluTyThrGlnAsnLySylsGluAglUlys 61
   |||||
DB 123 GATGAGCTTCACAGAGGACGACGAGATACCCAAACAAAGAGGAGACAGATC 182
   |||||
QY 62 IleYAsnLeuLeuTyraValIleYAsnLeuAlaIleLeuTyraAsnAsnGlnPhe 81
   |||||
DB 183 ATCAGAAACCTCATCAGACAGATCATCAGCTGGCCATCTTAATAGAAATATCAGTT 242
   |||||
QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
   |||||
DB 243 AATCAAGATAGCTAGATGATGAGAAATTTAAGAAAGTTCAATCAGCTTGCTATG 302
   |||||
QY 102 ThrValAlSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeu 121
   |||||
DB 303 ACCGTGTCAGTTTCCATCAGGTGATTAATACCTTTGACCGAATGTGTATCCAGGCTG 362
   |||||
QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
   |||||
DB 363 TTTAATCAATGACAGAGATGCTGCACCAATCATTCAGCGCACCTCATCTGCAGATCA 422
   |||||
QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
   |||||
DB 423 CAAAGACGGGTATATATGTGTGATCTTTTCAGATTGTGATTTTGGCTGCTGCTG 482
   |||||
QY 162 TyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLys 181
   |||||
DB 483 TATTAATCCTTTGGGATTTTAAACCCACTTACAAAACTATGTGTATCAACAA 542
   |||||
QY 182 MetLeuAspGluGluAsnIle 188
   |||||
DB 543 ATGTGTGATGAGAGAAACATA 563
   |||||
RESULT 15
BQ424670 896 bp mRNA linear EST 23-MAY-2002
LOCUS AGENCOURT 7833747 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6153217
DEFINITION 5', mRNA sequence.
ACCESSION BQ424670.1 GI:21119985
VERSION BQ424670.1 GI:21119985
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fda.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM13492 row: f column: 02
High quality sequence stop: 596.
Location/Qualifiers
1..896
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6153217"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:
Score:	6.15e-104	896
Percent Similarity:	947.00	186
Best Local Similarity:	98.94%	Conservative: 0
Query Match:	98.44%	Mismatches: 2
DB:	5	Indels: 0
		Gaps: 0

US-10-627-571-2 (1-188) x BQ424670 (1-896)

```

QY 1 MetAlAThAspVa1PheAsnSerLyAsnLeuAlaValGlnAglNylsYlLeu 20
   |||||
DB 97 ATGGCCACAGATGTCCTTAATCCAAAAACCTGGCCGCTCAGGACCAAAAGATCTTG 156
   |||||
QY 21 GLYMeTAspSerLySerLleAlaThrThLeuLeuAspPheSerSerGluVal 40
   |||||
DB 157 GGTAAATGGTGTCCAAATCATCGCACCACTTAATAGACACCAAGTATGAGTG 216
   |||||
QY 41 LeuAspGluLeuTyraValThrArgGluTyThrGlnAsnLysylsGluAglUlys 60
   |||||
DB 217 CTGATGAGCTTCACAGAGTACAGGAGTACACCAAAACAAAGAGGAGAGAG 276
   |||||
QY 61 LysIleYAsnLeuLeuTyraValIleYAsnLeuAlaIleLeuTyraAsnAsnGln 80
   |||||
DB 277 ATCATCAAGACCTCATCAAGACATCATCAGCTGCCATCTTAATAGAAATATCAG 336
   |||||
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
   |||||
DB 337 TTTAATCAAGATAGCTAGATGATGAGAAATTTAAGAGAAATTCATCAGCTTGCT 396
   |||||
QY 101 MetThValAlSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
   |||||
DB 397 ATGACCGGTGTCAGTTTCCATCAGGTGATTAATCCTTTGACCGAATGTGTATCAGG 456
   |||||
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
   |||||
DB 457 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCCACCTGCTGCCAG 516
   |||||
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
   |||||
DB 517 TCACATGACAGCGGTATTAATGTGTGATCATCTTTTCAGATTGTAAATTTTGGCTGCC 576
   |||||
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
   |||||
DB 577 TTGTATTAATCCTTTTNGAATTTTAAACCCACTTACAAAACTAGTATGTATCAAC 636
   |||||
QY 181 LysMetLeuAspGluGluAsnIle 188
   |||||
DB 637 AAAATGTGTGATGAGAGAAACATA 660
   |||||
RESULT 16
CN648259 998 bp mRNA linear EST 13-MAY-2004
LOCUS ILLUMIGN_MGC_30108 Katze_MMPB Macaca mulatta cDNA clone IBTMB:6795
DEFINITION 5' similar to Bases 72 to 998 highly similar to human TNFAIP8
(Hs.17839), mRNA sequence.
ACCESSION CN648259.1 GI:47161702
VERSION CN648259.1
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta

```


QY 22 LysMetValSerIysSerIleAlaThrThreuleIleAspAspThrSerSerGluValLeu 41
 DB 63 AAATGGTGTCCAAATCATCGCACCACTTATTAACACACAAATGGTGAGGCTG 122
 QY 42 AspGluLeuValThrArgGluValThrGlnAsnIleValGluValGly 61
 DB 123 GATGAGCTTACAGAGGACGACGAGGACTACCCAAACAGAGGAGGACGAGATC 182
 QY 62 IleLeuAsnLeuIleValThrValIleValLeuAlaIleLeuValArgAsnAsnGlnPhe 81
 DB 183 ATCAAGAACCTCATCAAGACATATCAAGCTGCGCATCTTATTAAGAAATATCACTT 242
 QY 82 AsnGlnAspGluLeuAlaLeuMetGluValPheValIleValValIleGlnLeuAlaMet 101
 DB 243 AATCAAGATGAGTATGATGATGAGAGAAATTTAAGAGAAAGTTCACTGAGCTGATG 302
 QY 102 ThrValValSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArgLeu 121
 DB 303 ACCGTGTCAGTTTCCATCAGTGCATTAATACCTTTACCGGAATGTGTTATCCAGGCTG 362
 QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaValSer 141
 DB 363 TTAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTGCAGATCA 422
 QY 142 HisGluValArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaVal 161
 DB 423 CATGACGGGTATTAATATGTTGATCATTTTCAATGTGTGAATTTTGGCTGCTTG 482
 QY 162 TyrAspProPheGluAsnPheValProHisLeuGlnIleValLeuValCysAspGluVal 181
 DB 483 TATTAATCTTTGGGAATTTTAACCCCACTTCAAAATCATGTGATGATCAACAA 542
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 543 ATGTTGATGATGAAGAACATA 563

RESULT 18 779 bp mRNA linear EST 12-JUL-2004
 LOCUS CR548728
 DEFINITION DKFZ469E1232 r1 469 (synonym: pklid) Pongo pygmaeus cDNA clone
 ACCESSION CR548728
 VERSION CR548728.1 GI:50242352
 KEYWORDS EST.
 SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 1 (bases 1 to 779) Mewes,H.W., Weill,B., Amid,C., Oeanger,A.,
 Bahr,A., Lauber,J., and Wiemann,S.
 Pongo pygmaeus mRNA (Bahr,A., Lauber,J., Mewes,H.W., et al.)
 Unpublished (2004)
 COMMENT Contact: MIPS
 MIPS Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
 This is the 5' sequence of the clone insert from S. Wiemann,
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);
 Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen
 (Hilden/Germany) within the cDNA sequencing consortium of the
 German Genome Project. This clone (DKFZp469E1232) is available at
 the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,
 Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:
 clone@rzpd.de Further information about the clone and the
 sequencing project is available at
 http://mips.gsf.de/projects/cdna/.
 Location/Qualifiers
 1..779
 /organism="Pongo pygmaeus"
 /mol_type="mRNA"
 /db_xref="taxon:9600"
 /clone="DKFZp469E1232"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,92e-102 Length: 779
 Score: 934.00 Matches: 183
 Percent Similarity: 98.94% Conservative: 3
 Best Local Similarity: 97.34% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 7 Gaps: 0
 US-10-627-571-2 (1-188) x CR548728 (1-779)

/tissue_type="kidney"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="469 (synonym: pklid)"
 /notes="Vector: pSport1_sfi; Site_1: SfiIA; Site_2: SfiIB"

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGluValIleLeu 20
 DB 184 GTGGCAGCATGTCCTTATTTCCAAACCTGCGCTTCAGGACCAAAAGAAATCTTG 243
 QY 21 GlyMetValSerIysSerIleAlaThrThreuleIleAspAspThrSerSerGluVal 40
 DB 244 GGTAAATGGTGTCCAAATCATCGCACCACTTATTAAGACGACGAGTATGAGGTG 303
 QY 41 LeuAspGluLeuValThrArgGluValThrGlnAsnIleValGluValGly 60
 DB 304 CTGGATGAGCTTACAGAGTACGACGAGGAGTACATCAAAACAAAGAGGACGAGAG 363
 QY 61 LysIleValAsnLeuIleValThrValIleValLeuAlaIleLeuValArgAsnGln 80
 DB 364 ATCATGAAGACCTCATCAAGACATCATCAAGCTGCGCATCTTATTAAGAAATATCAG 423
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluValPheValIleValValIleGlnLeuAla 100
 DB 424 TTTAATCAAGATGACGTATGATGATGAGGAAATTTAAGAGAAATTCATCAGCTTGCT 483
 QY 101 MetThrValValSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArg 120
 DB 484 ATGACCGTGTGATTTCCATCATCGTGCATTAATACCTTTGACCGGAATGTATCCAG 543
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaVal 140
 DB 544 CTGTTAATGAATGACAGAGATGCTGCACCAATCATTCGGGTGACCTCAGTCCAG 603
 QY 141 SerHisGluValArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaVal 160
 DB 604 TCACATGAGACGGTTAATATGCTTTGATCATTTTTCAGATTGTGATTTTGGCTGCC 663
 QY 161 LeuTyrAsnProPheGluAsnPheValProHisLeuGlnIleValCysAspGluVal 180
 DB 664 TTGTATTAATCTTTTGGGAATTTTAAACCACTTCAAAACCTAATGTGATGTATCAAC 723
 QY 181 LysMetLeuAspGluGluAsnIle 188
 DB 724 AAATGTGATGATGAAGAACATA 747

RESULT 19 876 bp mRNA linear EST 25-SEP-2001
 LOCUS B1752550
 DEFINITION 603021969P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5192561 5',
 mRNA sequence.
 ACCESSION B1752550
 VERSION B1752550.1 GI:15744128
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 876)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: c9abbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: LLM11481 row: j column: 18
 High quality sequence stop: 813.
 Location/Qualifiers

FEATURES

source

```

1..876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5192561"
/lab_host="NIH_MGC_114"
/clone_lib="NIH_MGC_114"
/notes="Organ: brain; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

```

ORIGIN

Alignment Scores:

Pred. No.:	2..96e-102	Length:	876
Score:	933.00	Matches:	186
Percent Similarity:	98.41%	Conservative:	0
Best Local Similarity:	98.41%	Mismatches:	2
Query Match:	96.99%	Indels:	1
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x B1752550 (1-876)

```

QY 1 MetAAThAaPValPheAaSerLyAaLeuAlaValGlnAglInLySlySleLeu 20
DB 118 ATGGCCACAGATGCTTTTAATCCAAAACCTGGCGGTCAGCACAAAAGATCTTG 177
QY 21 GILySMetValSerLySerLleAlaThrThrLeuLleAspAspThrSerSerGluVal 40
DB 178 GGTAAATGCTGTCAAATCCATCGCACCACTTAATAGACACAAAGTAGAGTGAGTG 237
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySlyGlnAglInLyS 60
DB 238 CTGGATGAGCTCTACAGAGTGACCAAGGAGTACACCAAAAACAGAGGAGCGAGAG 297
QY 61 LysIleLyAsnLeuLleLyThrValIleLySleuAlaIleLeuTyArgAsnAngln 80
DB 298 ATCATCAAAACCTCATCAAGACAGCATCAAGCTGGCAATCTTTATAGATATATCG 357
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLySlyPheLySlyValIleGlnLeuAla 100
DB 358 TTTAATCAAGATGAGCTAGCATTCAGAAATTTAAGAGAAAGTTCACTGCTGCT 417
QY 101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB 418 ATGACCGTGGTCAGTTTCATCATCGGTGATTATACCTTTGACCGGATGGTTATCCAG 477
QY 121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 478 CTGTAAATGAATGACGAGAGATGCTGCACCAATATTCAGCGCACCTCTGCAAG 537
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAlaIa 160
DB 538 TACATCGACGGGTTATTAATGTTGATCATTTTCAGATTGTAAATTTTGGGCTGCC 597
QY 161 LeuTyArgAsnPro-PheGlyAsnPhelysProHisLeuGlnLySlyLeuCyAspGlyLea 180

```

```

DB 598 TTGTAATATCCTTTTGGGATTTTAACCCCATTAACAAAATATGATGATATCA 657
QY 180 nlySMetLeuAspGluGlnAsnIle 188
DB 658 CAAAATGTTGATGAAGAGAACATA 682

```

RESULT 20

BP433798

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Sus scrofa (pig)
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Uenishi, H., Egunchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H., Okumura, N., Hamada, N., and Awata, T.
 PEDF (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries
 Nucleic Acids Res. 32 (1), D484-D488 (2004)
 Contact: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in Animal Genome Research Program (Japan) by National Institute of Agrobiological Sciences and STRF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA library
 Vector sequences were eliminated by RepeatMasker version 2002/07/13 and crossmatch version 0.990319
 Low quality bases were trimmed based on the quality values.

FEATURES

source

```

1..773
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="LNG010079A10"
/tissue_type="lung"
/dev_stage="adult"
/clone_lib="full-length enriched swine cDNA library, adult lung"

```

ORIGIN

Alignment Scores:

Pred. No.:	1..33e-101	Length:	773
Score:	927.00 <td>Matches:</td> <td>179</td>	Matches:	179
Percent Similarity:	98.94% <td>Conservative:</td> <td>7</td>	Conservative:	7
Best Local Similarity:	95.21% <td>Mismatches:</td> <td>2</td>	Mismatches:	2
Query Match:	96.36% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	5	Gaps:	0

US-10-627-571-2 (1-188) x BP433798 (1-773)

```

QY 1 MetAAThAaPValPheAaSerLyAaLeuAlaValGlnAglInLySlySleLeu 20
DB 154 GTGGCCACAGATGCTTTTAATCCAAAACCTGGCGGTCAGCACAAAAGATCTTG 213
QY 21 GILySMetValSerLySerLleAlaThrThrLeuLleAspAspThrSerSerGluVal 40
DB 214 GGTAAATGCTGTCAAATCCATCGCACCACTTATGATGAGACACAGATGCGAAGT 273
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySlyGlnAglInLyS 60
DB 274 CTGGATGAGCTCTACAGGCTGACCAAGAGTACACCAAGAACAGAGGAGCAGAGAA 333

```

QY 61 LysilelysaenleuilelythThValilelyleuAlalleleuTYrargAaenAgn 80
 DB 334 ATCATCAAAAACCTCATCAAAAACGATCAAGCTGGCCATTCTTCAGAAATATCA 393
 QY 81 PheasnlnaspglualeuAlaleuMeGluylusPhelyslyValHlsglnleuAla 100
 DB 394 TTTTACCAGATGAGCTGACATGATGAGAAATTTTAAAGAAAGTCCACCACTTGGCT 453
 QY 101 MetThValValserPheHlsglnValAspTYrThPheAspArgaenValleuSerArg 120
 DB 454 ATGACCGTGGTCACTTTTCATCAGTGGATTTTCACCTTGAACGGAATGCTATCCAGG 513
 QY 121 LeuleuAsngluCYsaArgluMeleuHlsglnleleleuHlsglnleuThrallys 140
 DB 514 CTGCTCATATGATGACAGAGCTGCTCCACAGATCATCCAGCTCATCCAGCCGACAG 573
 QY 141 SerHlsglYArGValAsnaenValPheAspHlserAspCySglupheleuAla 160
 DB 574 TCACACGACGAGATTAACAAGCTTTGATCATTTTCAAGATTGTGATTTCTTGGCCGCC 633
 QY 161 LeuTYrAsnProPheglYAsnPhelysProHlsglnleuCYsaAspGlylleasn 180
 DB 634 TTGTATATATCCCTTTGGAAATTTCAAAACCGACCTTACAGAACTGTGTATGTATCAAC 693
 QY 181 LysMeleuAspgluGluAsnle 188
 DB 694 AAAATGTTGGATGAGATTAACATA 717

RESULT 21
 CN789950

LOCUS 669 bp mRNA linear EST 26-MAY-2004
 DEFINITION 4124483 BARC 8BOV Bos taurus cDNA clone 8BOV_35M20 5', mRNA
 sequence.
 CN789950
 ACCESSION CN789950.1 GI:47685930
 VERSION CN789950.1
 KEYWORDS
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 669)
 AUTHORS Baumann, R.G., Baldwin, R.L., Sonstegard, T.S., Van Tassell, C.P. and Matukumalli, L.K.
 TITLE Construction and Analysis of a CDNA Library Generated From Intestinal Muscle and Epithelial Tissues of Holstein Cattle
 JOURNAL Unpublished (2004)
 COMMENT Contact: Richard G. Baumann
 Bovine Functional Genomics Lab
 ANRI
 BUDG 162: BARC-EAST, Beltsville, MD 20705, USA
 Tel: 3015048604
 Fax: 3015048744
 Email: rbaumann@anri.barc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt '- -trim fastaVector identified by cross match using options -mismatch 12 -mismatch 12
 plate: 35 row: M column: 20
 Seq primer: CCTATTAGGTGACACTATAGAAC
 High quality sequence stop: 669.
 Location/Qualifiers
 1. 669
 /organism="Bos taurus"
 /mol_type="mRNA"
 /strain="Holstein"
 /db_xref="taxon:9913"
 /clone="8BOV_35M20"
 /sex="Female"
 /tissue_type="Epithelial, Muscle"
 /dev_stage="Lactating, Neonatal"
 /lab_host="DH10B TONR"
 /clone_id="BARC 8BOV"

Alignment Scores:
 Pred. No.: 1,45e-101 Length: 669
 Score: 926.00 Matches: 179
 Percent Similarity: 98.40% Conservative: 6
 Best Local Similarity: 95.21% Mismatches: 3
 Query Match: 96.26% Indels: 0
 DB: 7 Gaps: 0

US-10-627-571-2 (1-188) x CN789950 (1-669)

QY 1 MetAlaThrAspValPheAsnSerlysaenleuAlaValGlnAgluyslyleu 20
 DB 90 GTGGCTACAGATGCTTTTAATTCAAAACCTGGCTTCAGACACAAAAGATCTTG 149
 QY 21 GlyMetValSerlyserlelealathThleuileasphrhrserSerGluVal 40
 DB 150 GGCAAAATGGCATCCAGTCCATGCAACCCCTCATGACACCAAGCATGAAGT 209
 QY 41 LeuAspGluLeuTYrArgValThrArgGluTYrThGlnAsnlysaenleuAlaGluys 60
 DB 210 CTGATGAGCTCTACAGGCTGACCAAGATACCCAGAACAAAAGAGCGAGAA 269
 QY 61 LysilelysaenleuilelythThValilelyleuAlalleleuTYrargAaenAgn 80
 DB 270 ATCATCAAGAACTCATCAAGACCGTCATCAACCTGCGCATCTTAAAGAAATTAAC 329
 QY 81 PheAsnGlnAspglualeuAlaleuMeGluylusPhelyslyValHlsglnleuAla 100
 DB 330 TTTTACCAGATGAGCTGACATGATGAGAAATTTTAAAGAAAGTTCATCAGCTTGGCT 389
 QY 101 MetThValValserPheHlsglnValAspTYrThPheAspArgaenValleuSerArg 120
 DB 390 ATGACCGTGGTCACTTTTCCACAGGTGACTTCACTTTGACCGGAACGTGTCTCAAG 449
 QY 121 LeuleuAsngluCYsaArgluMeleuHlsglnleleleuHlsglnleuThrallys 140
 DB 450 CTGTTAAATGAGTGAAGAGATGCTGACACAGATCATCCAGGCTACCTCACCAACAG 509
 QY 141 SerHlsglYArGValAsnaenValPheAspHlserAspCySglupheleuAla 160
 DB 510 TCACATGACGCTGTAAACAAGCTTTGATCACTTCTCAGATTGTGATTTCTTGGCGGCC 569
 QY 161 LeuTYrAsnProPheglYAsnPhelysProHlsglnleuCYsaAspGlylleasn 180
 DB 570 TTGTATATATCCCTTTGAAATTTTAAACCCACTTACAGAACTGTGTATGTATCAAC 629
 QY 181 LysMeleuAspgluGluAsnle 188
 DB 630 AAAATGTTGGATGAGATTAACATA 653

RESULT 22
 CA984413

LOCUS 942 bp mRNA linear EST 06-JAN-2003
 DEFINITION AGENCOURT 11295099 NIH-MGC 164 Mus musculus cDNA clone
 IMAGE:30146475 5', mRNA sequence.
 CA984413
 ACCESSION CA984413.1 GI:27517069
 VERSION CA984413.1
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 942)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/

/note="Organ: Intestine; Vector: PCMVSPORT6.1; Site_1: NCRI; Site_2: ESCR; Normalized cow cDNA intestinal mRNA library in PCMVSPORT6.1, constructed from equimolar mRNA pools derived from 5 sources, 4 lactating intestinal, 1 neonatal intestinal 4/5 lactating, proximal duodenum, jejunum, distal ileum, colon, 1/5 neonatal, proximal duodenum, jejunum, distal ileum"

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe and Dr. Mina
CDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAW062 row: d column: 04
High quality sequence start: 21
High quality sequence stop: 693.

FEATURES

source
1..942
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30146475"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_164"
/note="Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI;
Non-normalized full-length enriched library from pooled
mouse embryonic limb, maxilla and mandible, day 10.5 and
11.5 (size selected for the 0.5-1 kb fragments) Cloned
directionally, priming method: Oligo-dT. cDNA enrichment:
>1k bp, Average insert size 1.8k bp. Priming sequence:
5'GACCTAGTCTTAGATCGGACGGCGGCCCTT 3'. Tissue contributed
by, David Rowe. Library constructed by Reegen, Invitrogen
Corp."

ORIGIN

Alignment Scores:

Pred. No.: 4,01e-101 Length: 942
Score: 924.00 Matches: 178
Percent Similarity: 98.94% Conservative: 8
Best Local Similarity: 94.68% Mismatches: 2
Query Match: 96.05% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x CA984413 (1-942)

```

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
DB 176 ATGGCTACGAGTGTCTTCATTCCTCAAAACCTGGCCCTTCAGGCAAAAGAAAGATCTTG 235
QY 21 GilylMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 236 GGCAGAAATGGTATCCAAATCCATCCGACACGCTGATCGACGACACGACGACGAGTG 295
QY 41 LeuAspGluLeuThrArgValThrArgIuIyrrThrGlnAsnIysIleGluVal 60
DB 296 CTAGAGTACCTGTACAGGGTGAACCAAGAGTACACCAAGCAAGAGAGGCGGAGAG 355
QY 61 IysIleIysAsnLeuIleIysThrValIleIysLeuAlaIleuIyrrArgAsnAngin 80
DB 356 GTCATCAAGACCTCATCAAGAGGTCATCAAGCTGCGCTCTCCACAGAAACATTCAG 415
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluYsPheIysIysValHisGlnLeuAla 100
DB 416 TTCATCAAGACGAGCTGCGCTCATGAGAGATTCAGAGAGAGTGCACGACTTGGC 475
QY 101 MetThrValValSerPheHisGlnValAspIyrrThrPheAspArgAsnValIleuSerArg 120
DB 476 ATGACGGTGTGACGCTTCACCAAGTACACCTTCGACCGCAATGTGCTGTCCAG 535
QY 121 LeuLeuAsnGluCyArGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIys 140
DB 536 CTGGCTGAACGAGTGCAGAGAGCTCTCAACGAGATATTCAGGCGCACTTACCGCCAG 595
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160

```

```

DB 596 TCTCAGCAGACGGGTTAATATGTCCTTGGACCATTTTTCAGATTGTGATTTTGGCTGCC 655
QY 161 LeuTyraenProPhoGlyAsnPhelyProHisLeuGlnIysLeuCyAspGlyIleAsn 180
DB 656 TTGTCAATCCCTTTGGAAAAGTTTAAACCTCACTTACAGAAACCTTGGCAGCGCATCAAC 715
QY 181 IysMetLeuAspGluGluAsnIle 188
DB 716 AAAATGTTGGATGAAGAGACATA 739

```

RESULT 23
BI759027 893 bp mRNA linear EST 25-SEP-2001
LOCUS 603042814F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183355 5',
DEFINITION mRNA sequence.
ACCESSION BI759027.1 GI:15750605
VERSION BI759027.1 GI:15750605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11457 row: k column: 04
High quality sequence start: 5
High quality sequence stop: 869.

FEATURES

source

1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5183355"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 4.94e-101 Length: 893
Score: 923.00 Matches: 185
Percent Similarity: 98.94% Conservative: 1
Best Local Similarity: 98.40% Mismatches: 2
Query Match: 95.95% Indels: 1
DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x BI759027 (1-893)

```

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
DB 94 GTGGCCACAGATGTCTTAAATTCCTCAAAACCTGGCCCTTCAGGCAAAAGAAAGATCTTG 153

```

QY 21 G1yVmeValSerlySer1lea1aThrThreuleaSpaPThrSerGluVal 40
 Db 154 GGTAAATGGTGTCCAAATCCATGCGACCACTTATAGACGACAGAGTAGAGCTG 213
 QY 41 LeuApgGluLeuYrArGValThrArgGluYrThrGlnAnlySlySgluAgluLys 60
 Db 214 CTGATGAGCTCTACAGAGTAGACGAGGAGTACACCCAAACAGAGAGGAGGAGAG 273
 QY 61 LysIleYsaenleuIleYrThValIleYsleuA1leleuYrArGAsnArgGln 80
 Db 274 ATCATCAAGAACCTCATCAAGACGATCAAGCTGACCTTCTTATAGAAATATACG 333
 QY 81 PheAsnGlnApgGluLeuA1leuMetGluYrPheSlySlyValHisGlnLeuA1 100
 Db 334 TTTATATCAAGATGACGTACATTCATGAGAAATTTAGAGAAAGTTCTACCTTGCT 393
 QY 101 MetThrValValSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArg 120
 Db 394 ATGACCGTGTGATTCATTCAGTGTGATTAATACCTTTCACCGGATGTGTATCCAGG 453
 QY 121 LeuLeuAsnGluYrArGValMetLeuHisGlnIleleuGlnArgHisleuThrAlaLys 140
 Db 454 CTGTAAATGAAGGACAGAGAGTGTCCACCAATCATTCAGCGCCACCTCAGCCAG 513
 QY 141 SerHisGlyArGValAsnAsnValPheAspHisPheSerAspCysGluPheLeuA1Ala 160
 Db 514 TCCATGACGACGGGTTATATATGTTGTTGATCATTTTCAAGATTGTGAATTTG-GCTGCC 572
 QY 161 LeuTYrAsnProPheGlyAsnPhelysProHisleuGlnIleYsleuYrArGAsnGln 180
 Db 573 TTGTATATATCTTTTGGGAAATTTTAAACCCACTTCAAAAACATATGTGATGTAC 632
 QY 181 LysMetLeuApgGluGluAsnIle 188
 Db 633 AAAATGTTGATGAGAGAACATTA 656

RESULT 24
 BF607236 832 bp mRNA linear EST 01-APR-2001
 LOCUS MY1_000178 Mouse 9-day fetus cDNA library ICRFp522 Mus musculus
 DEFINITION cDNA clone ICRFp522C0838 5', mRNA sequence.
 ACCESSION BF607236
 VERSION BF607236.1 GI:13503728
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 832)
 Yabuywi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B. G.,
 Lehrsich, H. and O'Brien, J.
 Detection of a high number of novel genes in a 9-day mouse embryo
 cDNA library normalised by oligonucleotide fingerprinting
 Unpublished (2001)
 JOURNAL Contact: Hennig S
 COMMENT laboratory 123, dept. Lehrsich
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr. 63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1612
 Fax: +49 30 8413 1380
 Email: hennig@molgen.mpg.de
 ESTs are made from clones being representatives of clone clusters.
 Clone clusters were calculated from oligonucleotide fingerprints.
 PCR Primers
 FORWARD: 5'-GAGCTATTCAGAGTAGTA-3'
 BACKWARD: 5'-TATATGACTACTATAGGG-3'
 Seq primer: 5'-ATTAGGTGACACTATG-3'
 High quality sequence stop: 832.
 Location/Qualifiers
 1..832
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"

US-10-627-571-2 (1-188) x BF607236 (1-832)

ALIGNMENT Scores:
 Pred. No.: 1,04e-100 Length: 832
 Score: 920.00 Matches: 177
 Percent Similarity: 98.94% Conservative: 9
 Best Local Similarity: 94.15% Mismatches: 2
 Query Match: 95.63% Indels: 0
 DB: 2 Gaps: 0

ORIGIN
 /clone="ICRFp522C0838"
 /issue_type="whole_embryo"
 /dev_stage="embryonic 9-day"
 /lab_host="E.coli, XL1 blue"
 /clone_id="Mouse 9-day fetus cDNA library ICRFp522"
 /note="Vector: pSVSPORT1; Site 1: NotI; Site 2: SalI;
 library preparation by oligo dT priming of RNA. Clones can
 be ordered from the Resource Center in Berlin,
 http://www.rzpd.de."

QY 1 MetAlaThrAspValPheAsnSerlySAsnleuA1ValGlnA1aglnlySlyIleLeu 20
 Db 156 GTGGCTAAGATGCTTTCATATTCCAAAAACCTGGCGTTGAGGACAAAAGATCTCTG 215
 QY 21 G1yVmeValSerlySer1lea1aThrThreuleaSpaPThrSerGluVal 40
 Db 216 GGCMAAATGGTATCCAAATCCATGCGCACCCGCTGATCGACGACGACGAGCTG 275
 QY 41 LeuApgGluLeuYrArGValThrArgGluYrThrGlnAnlySlySgluAgluLys 60
 Db 276 CTGATGAGCTCTACAGAGTAGACGAGGAGTACACCCAGAAAGAGGAGGAGAGG 335
 QY 61 LysIleYsaenleuIleYrThValIleYsleuA1leleuYrArGAsnArgGln 180
 Db 336 GTCATCAAGAACTCATCAAGACGCTCATCAAGAGTTCAAGAAAGAGGTCACGCTTGC 395
 QY 81 PheAsnGlnApgGluLeuA1leuMetGluYrPheSlySlyValHisGlnLeuA1 100
 Db 396 TTCAATCAAGACGAGCTGCGCTCATGAGAGAGTTCAAGAAAGAGGTCACGCTTGC 455
 QY 101 MetThrValValSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArg 120
 Db 456 ATGACCGTGTGATTCATTCAGTGTGATTAATACCTTTCACCGGATGTGTATCCAGG 515
 QY 121 LeuLeuAsnGluYrArGValMetLeuHisGlnIleleuGlnArgHisleuThrAlaLys 140
 Db 516 CTGCTGAAGAGAGCGGAGAGCTCTCAACGAGATCATTCAGGCGCACCTTACCGCAGG 575
 QY 141 SerHisGlyArGValAsnAsnValPheAspHisPheSerAspCysGluPheLeuA1Ala 160
 Db 576 TCCACGAGGACGGGTTATATATGCTTTCACCATTTTTCAGATGTGATTTTGGCTGCC 635
 QY 161 LeuTYrAsnProPheGlyAsnPhelysProHisleuGlnIleYsleuYrArGAsnGln 180
 Db 636 TTGTACAAATCCCTTTGGAAGATTAAACCTTACAGAACTTTCAGAGGAGGAGGATCAAC 695
 QY 181 LysMetLeuApgGluGluAsnIle 188
 Db 696 AAAATGTTGATGAGAGAACATTA 719

RESULT 25
 CA469176 841 bp mRNA linear EST 09-MAR-2004
 LOCUS AGENCOURT 10735591 NIH MGC 154 Mus musculus cDNA clone
 DEFINITION IMAGE:10122433 5', mRNA sequence.
 ACCESSION CA469176
 VERSION CA469176.1 GI:24925528
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAgiLysIleLeu 20
 Db 95 GTGGCTACAGATGCTTCAATTCCTCAAAACCTGGCCCTTCAGGCAAAAGAAATCTGT 154
 QY 21 G1yMetValSerLysSerIleAlaThrThreutleAspAspThrSerSerGluVal 40
 Db 155 GGCAGAAATGGTATTCCTCAATTCATGCGCACCGCTGATCGACACACAGAGAGAGTG 214
 QY 41 LeuAspGluLeuValArgValThrArgGluThrGlnAsnLysGluAlaGluLys 60
 Db 215 CTAGATAGAGCTGTACAGGTGACCAAGAGATGACCCAGAACCAAGAGAGAGGCGGAGAG 274
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLysThrArgAsnGln 80
 Db 275 GTCATCAAGAACCTTCATCAAGACGCTCATCAAGTGGCCCTTCCTCCAGAACCAATCAG 334
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
 Db 335 TTCATCAAGACAGCTGGCGCTCATGAGAGATTCCAAGAGAGAGTTCACAGCTTGC 394
 QY 101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 395 ATGACGGTCTGCTAGCTTCCACAGGTAGAGACCTTCGACCGCATGTGCTGTCCAGG 454
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 455 CTGCTGAACAGAGTCCAGAGCTCTTCACAGATCTTCAGCGCCCTTACCGCCAG 514
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 515 TCTCAGCAGAGGTTAATTAATGCTTTGACCATTTTCAGATTGTGATTTTGGCTGCC 574
 QY 161 LeuTyrAsnProPheGlyAsnPhelyProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 575 TTGTACATCCCTTTGGAAAGTTTAAACCTCACCTTACAGAACTTTCCGACGGATCAC 634
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 635 AAAATGTTGATGAGAGACATTA 658

RESULT 27
 AL558873 559 bp mRNA linear EST 02-APR-2004
 LOCUS AL558873 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DJ015H04 5-PRIME, mRNA sequence.
 ACCESSION AL558873
 VERSION AL558873.3 GI:46184260
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 559) Gruber, C., Jessee, J. and Polayes, D.
 REFERENCE Full-length cDNA libraries and normalization
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Unpublished (2001)
 JOURNAL On Feb 15, 2001 this sequence version replaced gi:31283006.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3485.f

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0DJ015D02Q01&c=3485.f.

FEATURES

source 1..559
 location/Qualifiers
 /organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ015H04"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,07e-100 Length: 559
 Score: 918.00 Matches: 180
 Percent Similarity: 98.90% Conservative: 0
 Best Local Similarity: 98.90% Mismatches: 2
 Query Match: 95.43% Indels: 0
 DB: 1 Gaps: 0

US-10-627-571-2 (1-188) x AL558873 (1-559)
 QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAgiLysIleLeuGly 21
 Db 3 GCCACAGATGCTTAAATTCCTCAAAACCTGGCCCTTCAGGCAAAAGAAATCTGTGGGT 62
 QY 22 LysMetValSerLysSerIleAlaThrThreutleAspAspThrSerSerGluValLeu 41
 Db 63 AAAATGGTGTCCAAATTCATCGCCACACCTTAAATAGACACCAAGTATGAGTCTGT 122
 QY 42 AspGluLeuValArgValThrArgGluThrGlnAsnLysGluAlaGluLysLys 61
 Db 123 GATGACCTGTACAGAGTGCACAGGAGTACACCAAAACCAAGAGAGGAGAGAGATC 182
 QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLysThrArgAsnGlnPhe 81
 Db 183 ATCAAGAACTCTATCAAGACAGTCAACAGTCAACCTGCGCATTTTATGGAATTAATCAATT 242
 QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAlaMet 101
 Db 243 AATCAAGATGAGCTTATGATGAGAGAAATTTAAGAGAAATTCATCATGCTTGTCTATG 302
 QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
 Db 303 ACCGTGATGATTTTCATCATGAGTGAATTAACCTTGACCGGATGTATTCAGGCTG 362
 QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
 Db 363 TTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCACCTCATCTGCCAAGTCA 422
 QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleu 161
 Db 423 CATGACGGGTTTAAATTAATGTTGATTCATTTTCAGATTGTGAATTTTGGCTGCTGTG 482
 QY 162 TyrAsnProPheGlyAsnPhelyProHisLeuGlnLysLeuCysAspGlyIleAsnLys 181
 Db 483 TATTAATCCCTTTTGGAAATTTTAAACCCCACTTACAAAAATATAGTGTGKATTAACAA 542
 QY 182 MetLeu 183
 Db 543 ATGTTG 548
 RESULT 28
 BP302623 580 bp mRNA linear EST 17-SEP-2004
 LOCUS BP302623 Sugano cDNA library, macrophage Homo sapiens cDNA clone
 DEFINITION MPE06324, mRNA sequence.
 ACCESSION BP302623
 VERSION BP302623.1 GI:52231583
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 580)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.
 Location/Qualifiers

FEATURES

source
 1..580
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MPB06324"
 /cell_type="macrophage"
 /clone_lib="Sugano cDNA library, macrophage"

ORIGIN

Alignment Scores:
 Pred. No.: 1,12e-100 Length: 580
 Score: 918.00 Matches: 180
 Percent Similarity: 99.45% Conservative: 0
 Best Local Similarity: 99.45% Mismatches: 1
 Query Match: 95.43% Indels: 0
 Gaps: 0

US-10-627-571-2 (1-188) x BP302623 (1-580)

QY 8 SerLyAsnLeuAlaValGlnAglInLylsVleuGlySerMetValSerLysSer 27
 |||||
 DB 2 TCCAAAAACCTGGCCCTTCAGGACAAAAGAAATCTGGGTAAATGCTCCAAATCC 61
 |||||
 QY 28 IleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTyrArgVal 47
 |||||
 DB 62 ATCGCACACACTTAATAGACACACAAAGTAGAGAGTGTGATGAGCTTACAGAGTG 121
 |||||
 QY 48 ThrArgGluTyrThrGlnAsnLysGluAglLysLysIleLysAsnLeuIleLys 67
 |||||
 DB 122 ACCAGGAGTACACCCAAAACAAAGAGAGGACAGAGATCATCAAGAACTCATCAAG 181
 |||||
 QY 68 ThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnAspGluLeuAla 87
 |||||
 DB 182 ACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAGTTATATCAAGATGACTAGCA 241
 |||||
 QY 88 LeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrValValSerPheHis 107
 |||||
 DB 242 TTGATCGAAGAAATTTAGAGAAAGATTCATCAGCTTGTATGACCGTGTCACTTCCAT 301
 |||||
 QY 108 GlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGlu 127
 |||||
 DB 302 CAGGTGATTAATACCTTTCAGCCGGAATGTATTATCAGCGCTTAAATGATGACAGAG 361
 |||||
 QY 128 MetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgValAsnAsn 147
 |||||
 DB 362 ATGCTGCACCAATCATTCAGGCCACCTCATCGCCAAAGTCATCATGAGCGGGTTAATAT 421
 |||||
 QY 148 ValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTyrAsnPropheGlyAsn 167
 |||||
 DB 422 GTGTTTGCATCTTTTCAAGTTGTGAATTTTGGCTGCTGTATATCTTTTGGGAAT 481
 |||||
 QY 168 PheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeuAspGluGluAsn 187
 |||||
 DB 482 TTTAAACCCCATTAACAAAACATATGTATGATGATCAACAAAATGTTGATGAGAGAAC 541
 |||||
 QY 188 Ile 188
 |||||
 DB 542 ATA 544

RESULT 29
 BGI46795 651 bp mRNA linear EST 01-FEB-2001
 LOCUS
 DEFINITION mab95h09.y1 NCI-CGAP_Sp2 Mus musculus cDNA clone IMAGE:397813 5
 similar to TR:095379 095379 MDC-3.13 ISOFORM 2. [2] TR:09UP47 ;
 mRNA sequence.

ACCESSION BGI46795.1 GI:12650203
 VERSION BGI46795
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 651)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Email: cgapbs-remail.nih.gov
 Contact: Robert Strausberg, Ph.D.
 Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILM)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILM at:
<http://image.llnl.gov>

MGI:1478145
 Seg primer: -40RP from Gibco
 High quality sequence stop: 436.
 Location/Qualifiers

FEATURES

source

1..651
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:397813"
 /tissue_type="NK cells (flow-sorted)"
 /lab_note="DH10B (T1-resistant)"
 /clone_lib="NCI-CGAP_Sp2"
 /note="Organ: spleen; Vector: pCMV-SPORT6 (Life Technologies); mRNA made from flow-sorted NK cells, cDNA made by oligo-dT priming. Directionally cloned. Average insert size 1.5 kb. Primary library, non-amplified. cDNA library Preparation: David B. Krizman, Ph.D."

ORIGIN

Alignment Scores:
 Pred. No.: 2,28e-100 Length: 651
 Score: 916.00 Matches: 177
 Percent Similarity: 98.40% Conservative: 8
 Best Local Similarity: 94.15% Mismatches: 3
 Query Match: 95.22% Indels: 0
 Gaps: 0

US-10-627-571-2 (1-188) x BGI46795 (1-651)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAglInLysLysIleLeu 20
 |||||
 DB 88 ATGGCTACAGATGCTTCAATCCAAAAACCTGGCGCTTCAGGCACAAAGAAATCCCG 147
 |||||
 QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 |||||
 DB 148 GGCAAAAGATGATCCAAATCCATCCCAACAGCTGATGACAGACACCAAGCAGAGGG 207
 |||||
 QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGluAglLys 60
 |||||
 DB 208 CTAGATGAGCTTACAGGCTGACCAAGAGATACCCAGAACAAAGAGGCGGAGAGG 267
 |||||
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
 |||||
 DB 268 GTCATCAAGAACTCATCAAGAGCGGTCAATCAAGCTGCGCTCTCCACAGAAATCATAG 327

QY 81 PheanGlnAspGluLeuAlaLeuMetGlnLysPheLysValHisGlnLeuAla 100
 Db 328 TTCAATACAGACGAGCTGGCCCTCAATGAGAGAGTTCAAGAGAGAGGACACCTGGCC 387
 QY 101 MetThrValAlaSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 388 ATGACGGTGGTCACTTCCACACAGGTAGATGACCTTCCAGCCGCAATGGCTGTCCAGG 447
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 448 CTGCTGAACGAGTGCCTCAGAGCTCTACACGAGATCTTACGCCCACTTACCGCCAG 507
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 508 TCTACAGGACGGGTAAATTAATGCTGTGACCATTTTCAGATGTGATTTTGGCTGCC 567
 QY 161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 568 TTGTACATCCCTTTGAGAAAGTTTAAACCTCATTACAGAAACTTTCGACGCGATCAC 627
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 628 AAAATGTTAGTGAAGAGAACATA 651

RESULT 30
 AUI35377 816 bp mRNA linear EST 02-AUG-2002
 LOCUS AUI35377 PLACE1 Homo sapiens cDNA clone PLACE1001920 5', mRNA
 DEFINITION
 ACCESSION AUI35377
 VERSION AUI35377
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 816)
 AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
 TITLE HRI human cDNA project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomic@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
 FEATURES
 source
 1. 816
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1001920"
 /tissue_type="placenta"
 /clone_id="PLACE1"
 /note="Vector: pME18SFL3"

ORIGIN
 Alignment Scores:
 Pred. No.: 3,086-100 Length: 816
 Score: 916.00 Matches: 184
 Percent Similarity: 97.89% Conservative: 2
 Best Local Similarity: 96.84% Mismatches: 2
 Query Match: 95.22% Indels: 0
 DB: 1 Gaps: 2
 US-10-627-571-2 (1-188) x AUI35377 (1-816)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnIleGlnLysIleLeu 20
 Db 101 GTGGCCACAGATATCTTTAATTCAAAACCTGGCCGTTTCAGCACAAAAGAAATCTTG 160
 QY 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 161 GGTAAATGTGTCCAAATTCATCGCCACCACTTAAATAGACACACAGTAGTAGGCTG 220
 QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLys 60
 Db 221 CTGATGAGCTCTACAGAGTACAGGAGTATACCCAAACAAAGAGAGGACAGAGAG 280
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
 Db 281 ATCATCAAGAACCTCATCAAGACAGCATCAACCTGGCCATTTTAAATAGAAATTAAC 340
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGlnLysPheLysLysValHisGlnLeuAla 100
 Db 341 TTTAATCAAGATGAGCTAGCATGATGAGAGAAATTTAAGAGAAAGTTCAATGACTTGT 400
 QY 101 MetThrValAlaSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 401 ATGACCGTGGTCACTTCCATCAGGTGATTAACCTTGAACCGAATGTGTATCCAGG 460
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 461 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTCCAG 520
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 521 TCACATGAGACGGGTAAATTAATGCTGTGACCATTTTCAGATGTGAAATTTGGCTGCC 580
 QY 161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 581 TTGTATATACCTTTTGGAAATTTTAAACCCACTTCAAAAACCTAAGTAGATGATACAC 640
 QY 181 LysMet-LeuAspGluGluAsnIle 188
 Db 641 AAAATGTTGATGAAGAGAACATA 666

RESULT 31
 AK090316 1542 bp mRNA linear HTC 03-APR-2004
 LOCUS AK090316
 DEFINITION Mus musculus 14 days embryo lung cDNA, RIKEN full-length enriched library, clone:G630049H02 product:TNF-INDUCED PROTEIN G62-1 homolog (Homo sapiens), full insert sequence.
 ACCESSION AK090316
 VERSION AK090316.1 GI:26355502
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kikunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
(bases 1 to 1542)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hasegawa,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiroaka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,T., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kuribara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numata,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M., and Hayashizaki,Y.
Direct Submision
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
FEATURES
source
Location/Qualifiers
1.1542
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:G630049H02"
/db_xref="taxon:10090"
/clone="G630049H02"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="14 days embryo"
118..714
/note="unnamed protein product; TNF-INDUCED PROTEIN G32-1
homolog [Homo sapiens] (SPTRIQ9UP47, evidence: FASTV,
93.6%ID, 100%length, match=565)
putative"
/codon_start=1
/protein_id="BAC41167.1"
/db_xref="GI:26353503"
/translation="M:SEAEPEPVATDVNSKNLAVAQAKILGKMSKSIATLLID
DTSEVDELIVTVEYQNKKEAEVILNLIKTVLIRHNNQFNDLALMEKF
KKVQLQALVSPHOVEYTPDRNVLSTRLLKHLHLLIQRHLTAASHGRVNVFD
HSSDDFLAALNPFCKRPHQKLCIDGKIKMLDENT"
1521..1526
/note="putative"
polyA_signal

polyA_site 1542
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 1,26e-99 Length: 1542
Score: 914.00 Matches: 176
Percent Similarity: 98.40% Conservative: 9
Best Local Similarity: 93.62% Mismatches: 3
Query Match: 95.01% Indels: 0
DB: Gaps: 0
US-10-627-571-2 (1-188) x AK090316 (1-1542)
QY 1 MetAAThAAspValPheAsnSerIysAsnLeuAlaValAGlnAlaGlnIlySerIleLeu 20
:::|||||
DB 148 GGGCTACAGATGCTTCATTCACAAAACCTGGCGCTTCAGGCACAAAGAAATCTCG 207
QY 21 GILysMeValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
|||
DB 208 GGCAGAAATGATTCACAAATCCATCCGACCAAGCTGATGACGACACGACGACGAGTG 267
QY 41 LeuAspGluLeuTyArgValThrArgIuTyThrGlnAsnIlysgIuAgluIys 60
|||
DB 268 CTAGATGAGCTGTACAGGGGTGACCAAGGATGACCCGAGACAGAGAGGCGGAGAGG 327
QY 61 LysIleIysAsnLeuIleIysThrValIleIysLeuAlaIleLeuTyArgAsnEngin 80
|||
DB 328 GTCATCAAGAACCTCATCAAGAGGCTCATCAAGCTGGCGCTCTCCACAGAACATCAG 387
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIlyValIysGlnLeuAla 100
|||
DB 388 TTCATTCAGAGACGAGCTGGCGCTCATGAGAAAGTTCAAGAAAGATGACGACTTGC 447
QY 101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
|||
DB 448 ATGAGCGCTGCTGAGCTTCACACGAGTACGATTCACCTTCACCGCAATGTCTGTCCAG 507
QY 121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIys 140
|||
DB 508 CTGCTGAACGATGCGGACGAGGCTCTTACACGATCATTCACGCGACCTTACCGCAAG 567
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCyGluPheLeuAlaIa 160
|||
DB 568 TCTCAGCAGACGGGTTATATATGTCTTTCAGCATTTTCAGATTGTGATTTTGGCTGCG 627
QY 161 LeuTyArgAsnProPheGlyAsnPheIysProHisLeuGlnIlyLeuCyAspGlyIleAsn 180
|||
DB 628 TTGTGCAATCCCTTTGGAAAGTTTAAACCTCATTACAGAAACTTTGGGACGCGATCAA 687
QY 181 LysMetLeuAspGluGluAsnIle 188
|||
DB 688 AAAATGTTAGATGAAGAGAACATA 711
RESULT 32
BI688217 908 bp mRNA linear EST 18-SEP-2001
BI688217
LOCUS 603314718F1 NC1_CGAP_Mame Mus musculus cDNA clone IMAGE:5354452 5',
DEFINITION mRNA sequence.
ACCESSION BI688217
VERSION BI688217.1 GI:15650845
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at: <http://www.rockefeller.edu/imagelint/>

FEATURES	Location/Qualifiers
SOURCE	1. 908

ORIGIN

Alignment Scores:

Pred. No.:	1.08e-99	Length:	1908
Score:	912.00	Matches:	175
Percent Similarity:	98.40%	Conservative:	10
Best Local Similarity:	93.09%	Mismatches:	3
Query Match:	94.80%	Indels:	0
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x BI688217 (1-908)

Qy	1	MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu	20
Ds	85	ATGGCTACAGATGCTCTTCATTCCTCAAAAACCTGGCGCTTCAGCAGCAAAAGAACTCTG	144
Qy	21	GlyLysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerGluVal	40
Ds	145	GGCAAAATGGTATTCAAATTCATCCGCCACACAGCTGATGCAGACACGACGAGAGGTG	204
Qy	41	LeuAspGluLeuTyrArgValThrArgIuTyrThrGlnAsnLysLysGluAlaGluLys	60
Ds	205	CTGATGATGACCTGTACAGAGGTGACCAAGAGAGTACACCCAGAACAAAGAAAGAGCGGAGAGG	264
Qy	61	LysIleLysAsnLeuIleLysThrValIleLysIleAlaIleLeuTyrArgAsnAsnGln	80
Ds	265	GTCATCAAGAACTCTCATCAAGAGGTGCATCAAGCTGGCGCTTCACAGAACAAATCAG	324
Qy	81	PheAsnGlnAspGluLeuLysAlaLeuMetGlnLysPheLysValIleGlnLeuAla	100
Ds	325	TTCAATTCAGACGAGCTGGCGCTCATGAGAAATTCAGAAAGAGAGTCCACCACTGGCC	384
Qy	101	MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg	120
Ds	385	ATAGCGGTGCTGACTTCCACCAAGTGAAGATACACTTCGCAGCCGAAATGTCGTCCAGG	444
Qy	121	LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLysLeuThrAlaLys	140
Ds	445	CTCTCTAACAACAGTGGCCGAGAGCTCTCAACGAGATCATTTCAAGCGCCACTTACCGCCAG	504
Qy	141	SerHisGlyArgValAsnAsnValPheAspHisAspIleSerAspCysGlnPheLeuAla	160
Ds	505	TCTCAGCGAGCGGTTAAATATGCTTTTGAACATTTTCAGATGTGATTCCTTTGCGGCC	564
Qy	161	LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn	180
Ds	565	TTTGACATATCCCTTTGGAAAGTTAAACCTCACTTACAGAAACCTTGGACGCGCTCAC	624
Qy	181	LysMetLeuAspGlnGluAsnIle	188

Db 625 AAAATGTTGGATGAAGAGAACCTT 648

RESULT	33
LOCUS	B1683693
DEFINITION	B1683693 766 bp mRNA linear EST 18-SEP-2001 60330624.F1 NCI CGAP Mam6 Mus musculus cDNA clone IMAGE:5342123 5'
ACCESSION	B1683693
VERSION	B1683693
KEYWORDS	B1683693.1 GI:15646321
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

FEATURES

ORIGIN

Alignment Scores:

Pred. No.:	1,99e-99	Length:	766
Score:	909.00	Matches:	177
Percent Similarity:	98.41%	Conservative:	9
Best Local Similarity:	93.65%	Mismatches:	2
Query Match:	94.49%	Indels:	1
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x BI683693 (1-766)

QY	1	MeLaIaThrAspValPheAsnSerIlysAsnLeuAlaValGlnAlaGlnIleLysIleLeu	20
Db	142	ATGGCTACAGATGCTTCAATCCAAAACCTGGCCGCTTACAGCACAAAGAAATCTGT	20
QY	21	GIlySMetValSerIlySerIleAlaThrThreulIeaAspThrSerSerGluVal	40
Db	202	GGCAAAATGTATCCAAATCCATCCCAACAGCTGATCGACGACCCAGCAGGAGGTG	261
QY	41	LeuAspGluLeuTyArgValThrArgGluIuTyThrGlnAsnIleLysGluAlaGluLys	60
Db	262	CTAGATGAGCTTACAGGCTGACAGAGATGACCCGAAACAAAGAGAGGCGGAGAGG	322
QY	61	LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGln	80

Db 322 GTCATCAGAACTCATCAAGACGTCATCAAGCTGCGCTCTCCACAGAAACAATCAG 381
 QY PheAsnglAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 382 TTCATCAAGACGAGCTGCGCTCATGAGAACTTCAAGAGAGAGTGCACCGCTTGC 441
 QY 101 MetThrValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 442 ATGACGCTGCTGCTTCCACACGAGTAGAGTACACCTTCGACCGCATGTGCTCCAG 501
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 502 CTGCTGAACGAGGCTCGAGAGCTTCAACGAGATCATTCAGCGCACCTTACCGCCAG 561
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 562 TTCACGAGACGGGTAAATATATGCTTTCACCATTTTTCAGATTGTATTTTGGCTGCC 621
 QY 161 LeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnLysLeu-CysAspGlyIleAs 180
 Db 622 TTGTACATCCCTTTGGAAAGTTTAAACCTCACTTACAGAACTTTTGGCAGCGCATCA 681
 QY 180 nlySmcLeuAspGluGluAsnIle 188
 Db 682 CAAATGTTGGATGAAGAGAACTTA 706
 RESULT 34
 BI103517 772 bp mRNA linear EST 26-JUN-2001
 LOCUS BI103517 602889058F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044167
 DEFINITION 5', mRNA sequence.
 ACCESSION BI103517
 VERSION BI103517.1 GI:14554410
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 772)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL1121 row: C column: 16
 High quality sequence stop: 751.
 Location/Qualifiers
 1..772
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5044167"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1ib="NCI CGAP Kid14"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: Nci;
 Site 2: Salt; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ALIGNMENT Scores:

Pred. No.: 3,52e-99 Length: 772
 Score: 907.00 Matches: 177
 Percent Similarity: 98.41% Conservative: 9
 Best Local Similarity: 93.65% Mismatches: 2

ORIGIN

technology. Note: this is a NCI_CGAP Library."

Query Match: 94.28% Indels: 1
 DB: 4 Gaps: 0
 US-10-627-571-2 (1-188) x BI103517 (1-772)
 QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
 Db 86 GTGGCTACAGATGCTTCAATTCCAAAACCTGGCCGCTTCAGGCCAAAGAGATCTCG 145
 QY 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 146 GCAAAATGGATCCAAATCCATCCGACACGCTGATGACGACACCAAGAGAGG 205
 QY 41 LeuAspGluLeuTyrArgValThrArgLysTyrThrGlnAsnLysLysGluAlaGluLys 60
 Db 206 CTAGATGAGCTTACAGGCTGACCAAGAGTACACCAAGAACAAAGAGGCGGAGAGG 265
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
 Db 266 GTCATCAGAACTCATCAAGACGCTCATCAAGCTGCGCTCTCCACAGAACATCAG 325
 QY 81 PheAsnglAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 326 TTCATCAAGACGAGCTGCGCTCATGAGAACTTCAAGAGAGAGTGCACCGCTTGC 385
 QY 101 MetThrValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 386 ATGACGCTGCTGCTTCCACACGAGTAGAGTACACCTTCGACCGCATGTGCTCCAG 445
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 446 CTGCTGAACGAGTGCAGAGAGCTTCAACGAGATCATTCAGCGCACCTTACCGCCAG 505
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 506 TTCACGAGACGGGTAAATATATGCTTTCACCATTTTTCAGATTGTATTTTGGCTGCC 565
 QY 161 LeuTyrAsnProPheGly-AsnPhelysProHisLeuGlnLysLeu-CysAspGlyIleAs 180
 Db 566 TTGTACATCCCTTTGGCAAAGTTTAAACCTCACTTACAGAACTTTTGGCAGCGCATCA 625
 QY 180 nlySmcLeuAspGluGluAsnIle 188
 Db 626 CAAATGTTGGATGAAGAGAACTTA 650
 RESULT 35
 CB991862 785 bp mRNA linear EST 01-MAY-2003
 LOCUS AGENCOURT_13618130 NIH_MGC_148 Homo sapiens cDNA clone
 DEFINITION IMAGE:30337918 5', mRNA sequence.
 ACCESSION CB991862
 VERSION CB991862.1 GI:30286382
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 785)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Stefan Hanson
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM363 row: 1 column: 23
 High quality sequence stop: 629.

FEATURES
source

Location/Qualifiers
1. 785
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30337918"
/issue_type="pre-clampitic placenta"
/lab_host="DH10B Tona"
/clone_id="NH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTNN-3', size selected for average insert size 2.3 kb and normalized to R0.5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3.6e-99 Length: 785
Score: 907.00 Matches: 184
Percent Similarity: 97.87% Conservative: 0
Best Local Similarity: 97.87% Mismatches: 2
Query Match: 94.28% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x CB991862 (1-785)

QY 1 MetCAtArThrApAlPheAenSerLyAsnLeuAlValGlnAlGlnLySylLeu 20
DB 133 ATGCCCAAGATGCTTTAAATCCAAAACCTGGCCCTTACAGCACAAGAAAGATCTTG 192
QY 21 GLyLewMetValSerLySeriLeaIaThrTherLeuLeaSpThrSerSerGluVal 40
DB 193 GGTAAATGCTGCCAATCATCATGCCACCTTTATAGCAGACAGTAGTAGAGTG 252
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySylGluAlGluLyS 60
DB 253 CTGGATGAGCTTACAGAGTGACAGGAGTAGACCCAAAACAAGAGGAGGAGAG 312
QY 61 LySylLeuAsnLeuIleTyThrValIleLySleuAlIleLeuTyArgAsnArgGln 80
DB 313 ATCATCAAGAACTCATCAAGACAGCATCAAGCTGCCCATCTTTATAGGAATTAATCAG 372
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLySylPheLySylValIleGlnLeuAla 100
DB 373 TTTTATCAGATGAGCTTACATTCATGATGAGAAATTTAAGAAAGATTCAACAGCTTGCT 432
QY 101 MetThrValAlaSerPheHisGlnValAspTyThrPheAspArgAsnValIleSerArg 120
DB 433 ATGACCGTGTGCTGTTCCATCAGTGAGTTATACCTTTACCGGATGTGTTATCCAGG 492
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLyS 140
DB 493 CTCTTAAATGAATGCAGAGAGATGCTGCACCAATCATCTTACAGGCCCATCTCAGTCCAG 552
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
DB 553 TCCATAGGACGGGTATTAATGTCTTGATCATTTTCAATGTGTGAATTTTGGCTGCC 612
QY 161 LeuTyAsnProPheGlyAsnPheLyPheHisLeuGlnLySylCys-AspGlyIleAs 180
DB 613 TTGTATATATCTTTTGGGATTTTAAACCCACTTACAAAACATATGTGTATGTTATCA 672
QY 180 nLyMetLeu-AspGluGln 186
DB 673 CAAAATGTTGGATGAAGAG 692

RESULT 36
AK087478LOCUS
DEFINITION

AK087478 1836 bp. mRNA linear HTC 03-Apr-2004
Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30304C20 product:TNF-INDUCED PROTEIN GS2-1 homolog (Homo sapiens), full insert sequence.

ACCESSION

AK087478.1 GI:26104312

VERSION

AK087478.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

High-efficiency full-length cDNA cloning

TITLE

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

99279253

MEDLINE

10349636

PUBMED

2

REFERENCE

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

AUTHORS

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

TITLE

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

20499374

MEDLINE

11042159

PUBMED

3

REFERENCE

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.

AUTHORS

RIKEN integrated sequence analysis (RISA) system--384-format

TITLE

sequencing pipeline with 384 multicapillary sequencer

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE

11076861

PUBMED

4

REFERENCE

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

AUTHORS

Functional annotation of a full-length mouse cDNA collection

TITLE

Nature 409, 685-690 (2001)

JOURNAL

5

MEDLINE

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

AUTHORS

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

TITLE

Nature 420, 563-573 (2002)

JOURNAL

6 (bases 1 to 1836)

MEDLINE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murate, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

AUTHORS

Direct Submission

COMMENT

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Db 445 CTGTTAATGATGACAGAGATGCTGACCAATCATCTAGCCCGACCTACCTCCAG 504

Qy 141 SerHisgIyArGValAsnAsnValPheAspHisPheSerAspCySgIuPheLeuAla 160

Db 505 TCACATGACGCGGTGTTAATGATCTTGATCATTTTTCAGATTGGAATTTTGGCTGCC 564

Qy 161 LeuTyraNpRopHegIyAsnPhelyProHisLeuGlnIySLeuCyAspGlyTle 179

Db 565 TTGTAATATCCTTTTGGAAATTTTAAACCCCATTCACAAAACATATGATGATGCATC 621

RESULT 38

LOCUS B1687787 656 bp mRNA linear EST 18-SEP-2001

DEFINITION 603314107F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5354281 5', mRNA sequence.

ACCESSION B1687787

VERSION B1687787.1 GI:15650415

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 656)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <http://image.lnl.gov>
Plate: LMNL1900 row: m column: 02
High quality sequence, scop: 632.
Location/Qualifiers
1. 656
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5354281"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies, Inc. Investigator providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Alignment Scores:

Pred. No.: 1.14e-98 Length: 656

Score: 902.00 Matches: 173

Percent Similarity: 98.92% Conservative: 10

Best Local Similarity: 93.51% Mismatches: 2

Query Match: 93.76% Indels: 0

DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x B1687787 (1-656)

Qy 1 MetAlaThrAspValPheAsnSerIyAsnLeuAlaValGlnAlaGlnIySLeu 20

Db 87 GTGGCTACAGATGCTTCAATTCACAAAACCTGCGCTTCAGGCAAAAGAGATCTTG 146

Qy 21 G1YySeMetValSerIySerIleAlaThrThrLeuIleAspAspThrSerSergIuVal 40

Db 147 GGCAAAATGGTATTCAAATCATGCGCACCGTGTATTCACACACCAACGAGGTG 206

Qy 41 LeuAspGluLeuTyraGValThrArgGluTyThrGlnAsnIySgIuAlaGluIyS 60

Db 207 CTAGATAGAGGTGTAACAGAGGTGACCAAGAGTACACCAAGAAAGAGGCGAGAGG 266

Qy 61 LyeIleIyAsnLeuIleIySThrValIleIySLeuAlaIleLeuTyraGAsnGln 80

Db 267 GTCATCAAGAACTCATCAAGACGGTCATCAAGCTGCGCTCTCCACAGAACATACG 326

Qy 81 PheAsnGlnAspGluLeuAlaIleuMetGlnIySPhelySlySlySValIhIsgInLeuAla 100

Db 327 TTCATCAAGACGAGCTGCGCTCATGAGACAAAGTTCAAGAAAGAGTGCCACGCTTCC 386

Qy 101 MetThrValIyAsrPheHisIsgInValAspTyThrPheAspArgAsnValLeuSerArg 120

Db 387 ATGACGGTCTGTCAGCTTCCACAGGTAGAGTACACCTTCACCGCAATGTGCTGCAG 446

Qy 121 LeuLeuAsnGlnIySAsrArgIuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIyS 140

Db 447 CTGCTGAACGAGTGCAGAGAGCTCTTACACAGAGTCAATTACGCGCACCTTACCGCAAG 506

Qy 141 SerHisgIyArGValAsnAsnValPheAspHisPheSerAspCySgIuPheLeuAla 160

Db 507 TCACACGACGCGGTATATATGCTTGGACCATTTTTCAGATTGTGATTTTTCGCTGCC 566

Qy 161 LeuTyraNpRopHegIyAsnPhelyProHisLeuGlnIySLeuCyAspGlyTleAsn 180

Db 567 TTGTACATATCCTTTTGGAAAGTTTAAACCTCATTACAGAACTTGGCAGCATCAAC 626

Qy 181 LyeMetLeuAspGlu 185

Db 627 AAAATGTGTGATGAC 641

RESULT 39

LOCUS CR557522 707 bp mRNA linear EST 13-JUL-2004

DEFINITION DKFZp469E22119.F1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone

ACCESSION CR557522

VERSION CR557522.1 GI:50278801

KEYWORDS EST.

SOURCE Pongo pygmaeus (orangutan)

ORGANISM Pongo pygmaeus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo. 1 (bases 1 to 707)

AUTHORS Pousotka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R., Mewes, H.W., Weill, B., Amid, C., Osanger, A., Fodor, G., Han, M. and Wiemann, S.

TITLE Pongo pygmaeus mRNA (Pousotka, A., Albert, R., Moosmayer, P., et al.)

JOURNAL Unpublished (2004)

COMMENT Contact: MIPS

ORIGIN

Alignment Scores:

Pred. No.: 1. 707

Score: 102.00 Matches: 173

Percent Similarity: 98.92% Conservative: 10

Best Local Similarity: 93.51% Mismatches: 2

Query Match: 93.76% Indels: 0

DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x B1687787 (1-656)

Qy 1 MetAlaThrAspValPheAsnSerIyAsnLeuAlaValGlnAlaGlnIySLeu 20

Db 87 GTGGCTACAGATGCTTCAATTCACAAAACCTGCGCTTCAGGCAAAAGAGATCTTG 146

Qy 21 G1YySeMetValSerIySerIleAlaThrThrLeuIleAspAspThrSerSergIuVal 40

Db 147 GGCAAAATGGTATTCAAATCATGCGCACCGTGTATTCACACACCAACGAGGTG 206

Alignment Scores:

Pred. No.:	1 67e-98	Length:	707
Score:	901.00	Matches:	176
Percent Similarity:	99.44%	Conservative:	2
Best Local Similarity:	98.32%	Mismatches:	1
Query Match:	93.66%	Indels:	0
DB:	7	Gaps:	0

US-10-627-571-2 (1-188) x CRS57522 (1-707)

```

QY      1 MetAlaThrAspValPheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlySylleu 20
DB      170 GTGGCCACAGATGCTTTAATTCGAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTG 229
QY      21 GlyIyMetValSerIlySerIleAlaThrThleuIleAspAspThrSerSerGluVal 40
DB      230 GGTAATAATGGTGTCCAAATCCATCGCCACCACTTAATAGACACAGAGTACTGAGGTG 289
QY      41 LeuAspGluLeuTyArgValThrArgIlyTrhGlnAsnIlySylGlnAlaGluIly 60
DB      290 CTGGATGAGCTTCAAGAGTGCACCAAGGAGTACATCAAAAACAAAGAGGAGGAGAG 349
QY      61 IylIleIyAsnleuIleIySylThrValIleIySylleuAlaIleLeuTyArgAsnAngin 80
DB      350 ATCATCAAGAACTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGGAATATCAG 409
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluIySylPheIySylValHsglnLeuAla 100
DB      410 TTTAATCAAGATGAGTATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCT 469
QY      101 MetThrValAlaSerPheHsglnValAspTyTrhPheAspArgAsnValleuSerArg 120
DB      470 ATGACCGTGTGATGCTTTCATCAGTGCAGTATTAACCTTTCACCGGAATGTGTATCCAG 529
QY      121 LeuLeuAsnGluCysArgIleuMetLeuHsglnIleIleGlnArgHsisleuThralaIys 140
DB      530 CTGTTAATGAATGACAGAGAGTGCACCAAAATCATTCAGCGTCACTGCAGCAAG 589
QY      141 SerHsgIyArgValAsnAsnValPheAspHspSerAspCygluPheLeuAlaIa 160
DB      590 TCACATGAGCGGGTAAATATGCTTGTATCATTTTTCAGATTGTGATTTTGGCTGCC 649
QY      161 LeuTyAsnProPheGlyAsnPheIySylProHsisleuGlnIySylLeuCyAspGlyIle 179
DB      650 TTGTATTAATCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTAATGTGATGTATC 706

RESULT 40
LOCUS   CVO23086 566 bp mRNA linear EST 20-AUG-2004
DEFINITION 5463 Full Length cDNA from the Mammalian Gene Collection Homo sapiens cDNA 5' similar to BC007014, mRNA sequence.
ACCESSION CVO23086
VERSION   CVO23086.1 GI:51480836
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 566)
Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
Driscot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalan,P.O.,
Clingingsmith,T.R., Hartley,J.L., Esposito,D., Ched,D., Moore,T.,
Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
Vandenhaute,J., Cusick,M.E., Albaladejo,J.S., Hill,D.E. and Vidal,M.
Human ORFome Version 1.1: a Platform for Reverse Proteomics
Genome Res. (2004) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739

```

Email: Marc.Vidal@dfci.harvard.edu
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF results from a PCR reaction using an MGC full-length cDNA as template DNA and ORF specific primers
 PCR Primers
 FORWARD: ATGCACTCCGAGACAGAAATCCAA
 BACKWARD: CATATGTTCTCTCATCCACATTT
 Insert length: 566 Std Error: 21.00
 Plate: 11075 row: 05 column: F
 Seq primer: ACTGGCGGTGTTTACACGCTGACTGGAGAAAC
 High quality sequence start: 104
 High quality sequence stop: 565
 POLYA=No.

FEATURES

source	location/Qualifiers
1..566	organism="Homo sapiens"
	mol_type="mRNA"
	/db_xref="taxon:9606"
	/cisue_type="mixed"
	/clone_lib="Full Length cDNA from the Mammalian Gene Collection"
	/note="Vector: mixed. The ORFs were PCR amplified from the MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational Gateway cloning into pDONR223 donor vector. Reference: MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-Length Human and Mouse cDNA Sequences. PNAS, 2002, 99(126), 16899-16903"

ORIGIN

Alignment Scores:

Pred. No.:	1 64e-98	Length:	566
Score:	900.00	Matches:	176
Percent Similarity:	99.44%	Conservative:	1
Best Local Similarity:	98.88%	Mismatches:	1
Query Match:	93.56%	Indels:	0
DB:	7	Gaps:	0

US-10-627-571-2 (1-188) x CVO23086 (1-566)

```

QY      1 MetAlaThrAspValPheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlySylleu 20
DB      31 GTGGCCACAGATGCTTTAATTCGAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTG 90
QY      21 GlyIyMetValSerIlySerIleAlaThrThleuIleAspAspThrSerSerGluVal 40
DB      91 GGTAATAATGGTGTCCAAATCCATCGCCACCACTTAATAGACACAGAGTACTGAGGTG 150
QY      41 LeuAspGluLeuTyArgValThrArgIlyTrhGlnAsnIlySylGlnAlaGluIly 60
DB      151 CTGGATGAGCTTCAAGAGTGCACCAAGGAGTACATCAAAAACAAAGAGGAGGAGAG 210
QY      61 IylIleIyAsnleuIleIySylThrValIleIySylleuAlaIleLeuTyArgAsnAngin 80
DB      211 ATCATCAAGAACTCATCAAGACAGTCATCAAGCTGCAGCATCTTTATAGGAATATCAG 270
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluIySylPheIySylValHsglnLeuAla 100
DB      271 TTTAATCAAGATGAGTATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCT 330
QY      101 MetThrValAlaSerPheHsglnValAspTyTrhPheAspArgAsnValleuSerArg 120
DB      331 ATGACCGTGTGATGCTTTCATCAGTGCAGTATTAACCTTTCACCGGAATGTGTATCCAG 390
QY      121 LeuLeuAsnGluCysArgIleuMetLeuHsglnIleIleGlnArgHsisleuThralaIys 140
DB      391 CTGTTAATGAATGACAGAGAGTGTGCACCAAAATCATTCAGCGGCACCTCACTGCCAAG 450
QY      141 SerHsgIyArgValAsnAsnValPheAspHspSerAspCygluPheLeuAlaIa 160
DB      451 TCACATGAGCGGGTAAATATGCTTGTATCATTTTTCAGATTGTGAAATTTTGGCTGCC 510
QY      161 LeuTyAsnProPheGlyAsnPheIySylProHsisleuGlnIlySylLeuCyAspGlyIle 178

```


Db 511 TTGATATCTCTTTGGGAATTTAAACCCACCTTACAAACCTATGTGATGCT 564

RESULT 41
AL542926 779 bp mRNA linear EST 24-MAR-2004
LOCUS AL542926 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DB013Y017
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL542926
VERSION AL542926.3 GI:45718495
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 779)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:30548565.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 3485.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdnas=CS0DB013CE09QPl&c=3485.f.
Location/Qualifiers

FEATURES
source
1. 779
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB013Y017"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Alignment Scores:

Pred. No.: 4.38e-98 Length: 779
Score: 898.00 Matches: 178
Percent Similarity: 95.19% Conservative: 9
Best Local Similarity: 95.19% Mismatches: 0
Query Match: 93.35% Indels: 0
DB: 1 Gaps: 0

US-10-627-571-2 (1-188) x AL542926 (1-779)

Qy 2 AAlathzspValPheanSerLysasnLeuAlaValGlnAlaGlnLysLysIleLeuGly 21
Db 3 GCCACAGATGCTCTTAATTCACAAAACCTGGCCCTTAGGCACAAAAGATCTTGAGT 62
Qy 22 LysMetValSerLysSerIleLeuAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
Db 63 AAAMAGTGTTCCAATCATCGCCACCACTTAAMAMAMGACCAAGTAGTGAGTCTG 122
Qy 42 AspGluLeuLysThrValThrArgGluThrThrGlnAsnLysGluValGluValLys 61
Db 123 GATGAGCTCTACAGAGTGAACGAGGAGTACCCCAAAACAAAGAGGAGGAGCAATC 182
Qy 62 ILeYAsnLeuLysIleThrValIleLysLeuAlaIleLeuLysThrArgAsnAsnGlnPhe 81
Db 183 ATCAAGAACTCATCAAGACGATCATCAAGCTGCGCATCTTTATAGAAATAMCAGTTT 242

Qy 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
Db 243 AATCAAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 302
Qy 102 ThrValValSerPheHisGlnValAspThrThrPheAspArgAsnValLeuSerArgLeu 121
Db 303 ACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 362
Qy 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
Db 363 TTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
Qy 142 HisGluValArgValAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
Db 423 CATGACGGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 482
Qy 162 TyAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGluLysLeu 181
Db 483 TATATCTCTTTGGGAATTTTAAACCCACCTTACAAACCTATGTGATGATGATGATG 542
Qy 182 MetLeuAspGluLeuLeu 188
Db 543 ATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 563

RESULT 42
CD520304 747 bp mRNA linear EST 06-JUN-2003
LOCUS CD520304
DEFINITION IMAGE:30410510 5', mRNA sequence.
ACCESSION CD520304
VERSION CD520304.1 GI:31452022
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 747)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: NDCM200 row: m column: 15
High quality sequence stop: 499.
Location/Qualifiers

FEATURES
source
1. 747
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30410510"
/tissue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 191"
/note="Vector: pDON-LIB; Site 1: SfiI (ggccatcatggcc); Site 2: SfiI (ggccgcctggcc); Library is oligo-dT primed and directionally cloned. PBMC Peripheral Blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA adn Ionomycin. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGGCCGACGATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0.kb). 15/15 colonies contained inserts

ORIGIN

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

Alignment Scores:

Pred. No.:	1,67e-97	Length:	747
Score:	893.00	Matches:	176
Percent Similarity:	97.27%	Conservative:	2
Best Local Similarity:	96.17%	Mismatches:	5
Query Match:	92.83%	Indels:	0
DB:	6	Gaps:	0

US-10-6227-571-2 (1-188) x CD520304 (1-747)

```

QY      1 MetAlThAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLySylleu 20
DB      106 GTGGCCACGATGTCCTTAATTCAGAAAACCTGGCGCTCAGCAGCAAAAGATCTTG 165
QY      21 GlyLysMetValSerLySerLealAlaThrThrLeuLeuAspThrSerSerGluVal 40
DB      166 GGTAAATGGTGCCAAATCCATCGCCACCACTTATAGACGACCAACTATGAGTG 225
QY      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySylGluAlaGlu 60
DB      226 CTGGATGAGCTCTACAGAGTACAGCAGGAGTACACCAAAACAAAGAGCAGAGAG 285
QY      61 LysAlleLysAsnLeuLeuTyThrValleLysLeuAlaLeuLeuTyArgAsnAsn 80
DB      286 ATCATCAAGACCTCATCAAGACGATCATCAAGTGGCCATCTTATAGAAATATCAG 345
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB      346 TTTAATCAAGATGAGTATGATGATGAGAAATTTAAGAAAGTATCATGCTTGCT 405
QY      101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSer 120
DB      406 ATACCGTGGTCACTTCCATCATGAGTGGATTAATACCTTTACCGGAATGATGATCAG 465
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnleleGlnArgHisleuThrAla 140
DB      466 CTGTAAATGATATCGACAGAGAGTGGCACCAATCTTAAGGCCACCTCATGCGCAG 525
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB      526 TCACATGAGACGGGTATATATATGATGATCAATTTTCAATTTTGGCTGCC 585
QY      161 LeuTyArgAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCysAspGly 180
DB      586 TTGTATATATCCCTTTTGGGAATTTTAACCCCACTTCAAAAACATATGATGATGCA 645
QY      181 LysMetLeu 183
DB      646 CAAATGTTG 654

```

RESULT 43 BY743355 671 bp mRNA linear EST 17-DEC-2002
 LOCUS BY743355 RIKEN full-length enriched, bone marrow macrophage Mus
 DEFINITION musculus cDNA clone I830007M06 5', mRNA sequence.
 ACCESSION BY743355
 VERSION BY743355.1 GI:27169364
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 671)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schumacher, C., Gojodori, T., Balderelli, R., Hill, D. P., Bull, C.,
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Digrani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Q.D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takemura, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wyshew-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Zimnci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Waki, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Science Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
 Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, F.,
 Imotani, K., Ishii, Y., Itoh, M., Kasawa, I., Kawai, J., Kojima, Y.,
 Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
 Nomura, K., Numazaki, R., Ohno, M., Osato, N., Saito, R., Sakazume, N.,
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
 Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submissions
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichannel sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by David A. Hume (Dept. of Biochemistry
 and Microbiology/Parasitology Institute for Molecular Bioscience
 University of Queensland Brisbane, Q 4072 Australia) whose
 assistance we gratefully acknowledge.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES

source

Location/Qualifiers
 1..671
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="I830007M06"
 /tissue_type="bone marrow"

ORIGIN /cell_type="macrophage"
/clone.lib="RIKEN full-length enriched, bone marrow
macrophage"

Alignment Scores:

Pred. No.:	1,91e-97	Length:	671
Score:	892.00	Matches:	172
Percent Similarity:	97.31%	Conservative:	9
Best Local Similarity:	92.47%	Mismatches:	5
Query Match:	92.72%	Indels:	0
DB:	6	Gaps:	0

US-10-627-571-2 (1-188) x BF743355 (1-671)

```

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlyslIleu 20
DB 110 GTGGCTACAGATGTCTTCATTCCTCAAAACCTGGCCGTTCCAGCACAAAGAAAGATCTCG 169
QY 21 GlyValMetValSerIysSerIleAlaThrIleuIleAspAspThrSerSerGluVal 40
DB 170 GGCAAAATGGTATCCAAATCCATCCGACCCGCTGATCGACACACGACGAGGCTG 229
QY 41 LeuAspGluLeuTyraArgValThrArgGluTyrrThrGlnAsnIlyslValIleGluVal 60
DB 230 CTGATGAGCTGTACAGAGGTGACCAAGAGTACACCCAGAACAAAGAGGCGGAGAGG 289
QY 61 LysIleLysAsnLeuIleIysThrValIleLysLeuAlaIleLeuTyraArgAsnAngin 80
DB 290 GTGCTACAGAACTCTATCAAGACGCTCATCAAGCTGGCCGCTCTCCAGGAAACATCG 349
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 350 TTCATCAACAGACGAGCTGGCTCATCGAGAGAGTCAAGATTAAGGTGACGCTTGGCC 409
QY 101 MetThrValIalSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
DB 410 ATGACGGTGTGCTGCTTCCACCGAGTACAGTACCTTCCACCGCAATGTGCTGCAG 469
QY 121 LeuLeuAsnGluCyArgIleuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 470 CTGCTACAGAGTGGCGGAGGCTCTACACGAGATCATTCAGCGCACCTTACCGGCAAG 529
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
DB 530 TCTCAGGAGCGGCTTATATGCTTTCGACATTTTTCAGATTGTGATTMTTGGCTGCC 589
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnIlyslLeuCyAspGlyIleAsn 180
DB 590 TTGTACAAATCCCTTTGGAAAGTTTAAACCTCACTTACAGAACTTTGCGACGGCATCAAC 649
QY 181 LysMetLeuAspGluGlu 186
DB 650 AATAATGTTAGTCAAGAG 667

```

RESULT 44

BF140519 739 bp mRNA linear EST 24-OCT-2000
LOCUS 601787536F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4015160 5',
DEFINITION mRNA sequence.

ACCESSION BF140519
VERSION BF140519.1 GI:10979546
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 739)

REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov

FEATURES

source

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
Plate: LLM9260 row: h column: 09
High quality sequence stop: 709.
Location/Qualifiers

1..739

/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH 11"
/db_xref="taxon:10090"
/clone="IMAGE:4015160"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone.lib="NCI_CGAP_Lu30"
/note="Organ: lung; Vector: pCMV-SPORT6; Site:1: Nci1;
Site:2: Sail; transgenic model WNT-1, expression driven by
MMTV-ltr enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:

Pred. No.:	2.88e-97	Length:	739
Score:	891.00	Matches:	176
Percent Similarity:	98.40%	Conservative:	9
Best Local Similarity:	93.62%	Mismatches:	3
Query Match:	92.62%	Indels:	1
DB:	2	Gaps:	0

US-10-627-571-2 (1-188) x BF140519 (1-739)

```

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlyslIleu 20
DB 100 GTGGCTACAGATGTCTTCATTCCTCAAAACCTGGCCGTTCCAGCACAAAGAAAGATCTCG 159
QY 21 GlyValMetValSerIysSerIleAlaThrIleuIleAspAspThrSerSerGluVal 40
DB 160 GGCAAAATGGTATCCAAATCCATCCGACCGCTGATCGACACGACGAGGCTG 219
QY 41 LeuAspGluLeuTyraArgValThrArgGluTyrrThrGlnAsnIlyslValIleGluVal 60
DB 220 CTGATGAGCTGTACAGGCTGACCAAGAGTACACCCAGAACAAAGAGGCGGAGAG 279
QY 61 LysIleLysAsnLeuIleIysThrValIleLysLeuAlaIleLeuTyraArgAsnAngin 80
DB 280 GTCATCAAGAACTCTATCAAGACGCTGACAGCTGGCCGCTCTCCACGAGAAATCAG 339
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 340 TTCATCAACAGACGAGCTGGCTCATGAGAAAGTTCAAGAAAGGTGACCAAGCTTGGC 399
QY 101 MetThrValIalSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
DB 400 ATGACGGTGTGCTGCTTCCACGAGTACAGTACACTTTCAGACCGCACTGCTGCAG 459
QY 121 LeuLeuAsnGluCyArgIleuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 460 CTGCTACAGAGTGGCGGAGGCTCTTACACGAGATCATTCAGCGCACCTTACCGCAAG 519
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
DB 520 TCTCAGGAGCGGCTTATATGCTTTCGACATTTTTCAGATGTGATTMTTGGCTGCC 578
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnIlyslLeuCyAspGlyIleAsn 180
DB 579 TTGTACAAATCCCTTTGGAAAGTTTAAACCTCACTTACAGAACTTTGCGACGGCATCAAC 638
QY 181 LysMetLeuAspGluGlu 186

```

Db 639 AAAATGTCGATGACAGAAACATA 662

RESULT 45
CR771358

LOCUS
DEFINITION DKFZp46910535_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
CR771358
ACCESSION DKFZp46910535 5', mRNA sequence.
VERSION CR771358.1 GI:52614631
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE 1 (bases 1 to 692)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Weilenreuther,R.,
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
Pongo pygmaeus mRNA (Poustka,A., Albert,R., Moosmayer,P., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for
ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp46910535
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source 1..692
location/Qualifiers

/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp46910535"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/note="Vector: pSport1_sfl; Site_1: sfla; Site_2: sflb"

ORIGIN

Alignment Scores:

Pred. No.: 4,6e-97 Length: 692
Score: 889.00 Matches: 178
Percent Similarity: 98.90% Conservative: 2
Best Local Similarity: 97.80% Mismatches: 2
Query Match: 92.41% Indels: 1
DB: 7 Gaps: 0

US-10-627-571-2 (1-188) x CR771358 (1-692)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysLysIleLeu 20
Db 148 GTGGCCACAGATGCTTTAAATTCCAAACCTGCGCTTCAGGACAAAGAAATCTTG 207
QY 21 GlyLysMetValSerIysSerIleAlaThrLeuIleAspAspThrSerSerGluVal 40
Db 208 GGTAAGATGTGTCCAAATCCATCGCCACACCTTAATAGACGACAGAGTAGAGTG 267
QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLys 60
Db 268 CTGGATGAGCTCTCAAGAGTGAACGAGGAGTACATCAAAACAAAGAGGAGGAGAG 327
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
Db 328 ATCATCAAGAACCTCATCAAGACATCAAGCTGCGCATTTCTTAAGAAATATCAAG 387
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db 388 TTAAATCAAGATGAGCTACATGATGAGAAATTAAGAAAGTTCATCAGCTTGCT 447

QY 101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
Db 448 ATGACCGTGTGATGATTCATCAGATGATTAATACCTTGACCGGAATGTGTTATCCAG 507
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db 508 CTGTTAAATGATGACGAGAGATGCTGCACCAATCATTCACGCTCACCCTCAGCCAG 567
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
Db 568 TCACATGACGCGGTTAAATATGTCTTTCATCATTTTCAGATTCGATTTTGGCTGCC 627
QY 161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleLeu 180
Db 628 TTGTAAATCTTTTGGGATTTTAAACCCACCTTACA-AAACTAATGATGATGATCAAC 686
QY 181 LysMet 182
Db 687 AAAATG 692

Search completed: July 28, 2005, 21:30:47
Job time : 3024 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: July 28, 2005, 15:10:58 ; Search time 3088 Seconds
(without alignments)
9430.565 Million cell updates/sec

Title: US-10-627-571-1_COPY_100_700

Perfect score: 601

Sequence: 1 cctcagcgccgcgcgcgt.....tggatgaagagaacatatga 601

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_srs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	588.4	97.9	1892	9 AF070671	Homo sapi
2	588.2	97.9	1943	6 BD153536	human
3	586.8	97.6	1921	6 BD156880	Primer fo
4	586.8	97.6	1921	6 AX877846	Sequence
5	586.8	97.6	1921	6 AK001931	Homo sapi
6	564.4	93.9	1729	6 BD156785	Primer fo
7	564.4	93.9	1729	6 AX877673	Sequence
8	564.4	93.9	1729	9 AK001850	Homo sapi
9	564.4	93.9	1814	9 BC005352	Homo sapi
10	564.4	93.9	1915	9 AF099935	Homo sapi
11	564.4	93.9	1988	6 C0726075	Sequence
12	564.4	93.9	2003	9 AF099936	Homo sapi
13	564.4	93.9	2081	9 BC007014	Homo sapi
14	564.4	93.9	2502	9 AK097284	Homo sapi
15	564.4	93.9	112626	9 AC035144	Homo sapi
16	564.4	93.9	156277	9 AC026795	Homo sapi
17	564.4	93.9	158057	9 AC138612	Homo sapi
18	561.2	93.4	573	9 CR457137	Homo sapi
19	547.6	91.1	816	6 BD149395	Primer fo

20	547.6	91.1	816	6 AX869333	Sequence
21	500.8	83.3	1766	9 AK097884	Homo sapi
22	454.2	75.6	237561	2 AC095257	Rattus no
23	444.4	73.9	1728	10 BC009090	Mus muscu
24	444.4	73.9	184327	10 AC120859	Mus muscu
25	408.6	68.0	2009	5 AJ720906	Gallus ga
26	394.6	65.7	587	6 BD149633	Primer fo
27	394.6	65.7	587	6 AX869571	Sequence
28	355.6	59.2	1559	5 CR760636	Xenopus t
29	354	58.9	1193	5 BC072904	Xenopus l
30	304	50.6	515	9 AF098933	Homo sapi
31	257.4	42.8	164258	2 AC118938	Mus muscu
32	257.4	42.8	195574	2 AC107848	Mus muscu
33	256.4	42.7	3986	5 BC052765	Danio rer
34	256.4	42.7	205949	5 BX649252	Zebrafish
35	255.8	42.6	188937	2 AC012678	Homo sapi
36	255.8	42.6	189796	9 AC073964	Homo sapi
37	254.4	42.3	1248	6 C0736656	Sequence
38	254.4	42.3	1986	6 C0841622	Sequence
39	254.4	42.3	1986	9 AK123281	Homo sapi
40	252.4	42.0	1589	5 AJ720336	Gallus ga
41	250.2	41.6	110000	2 AC115187_1	Continuation (2 of
42	250.2	41.6	256285	2 AC115505	Rattus no
43	250.2	41.6	273325	2 AC129440	Rattus no
44	231.2	38.5	1938	5 BC076797	Xenopus l
45	225.4	37.5	1813	5 BC053167	Danio rer
46	225.4	37.5	189797	5 BX927313	Zebrafish
47	220.6	36.7	134497	2 CR407586	Danio rer
48	213.8	35.6	2175	5 BC061657	Xenopus l
49	209.6	34.9	1357	5 BC053238	Danio rer
50	207.8	34.6	32360	9 AC005339	Homo sapi
51	204.2	34.0	2209	9 BC017672	Homo sapi
52	203.4	33.8	107568	2 AC143398	Rattus no
53	203.4	33.8	230310	2 AC120079	Rattus no
54	199.8	33.2	2307	10 BC032199	Mus muscu
55	199.8	33.2	180238	2 AC073688	Mus muscu
56	199.8	33.2	222610	2 AC026385	Mus muscu
57	199.8	33.2	232190	2 AC074167	Mus muscu
58	199.6	33.2	561	6 C0732643	Sequence
59	195.4	32.5	149954	2 AC151645	Dasyus n
60	182	30.3	1200	9 BC063014	Homo sapi
61	182	30.3	1268	6 AX078255	Sequence
62	182	30.3	3116	9 HSM804686	Sequence
63	182	30.3	154526	2 AL592111	Homo sapi
64	182	30.3	159148	2 AL592424	Human DNA
65	180.4	30.0	1196	9 AK027120	Homo sapi
66	180.4	30.0	1433	10 BC079019	Rattus no
67	180.4	30.0	174234	10 AC117098	Rattus no
68	179.4	29.9	1198	10 BC058679	Mus muscu
69	179.4	29.9	148675	10 AC131769	Mus muscu
70	179.4	29.9	205702	10 AC140190	Mus muscu
71	179.4	29.9	206230	10 AC084272	Mus muscu
72	178.8	29.8	207952	2 AC151020	Callithri
73	176	29.3	555	5 CR457375	Homo sapi
74	174	29.0	181230	2 AC149856	Papio anu
75	174	29.0	182245	2 AC149842	Papio anu
76	171.6	28.6	1171	9 AF271774	Homo sapi
77	170.4	28.4	208275	2 AC147191	Orculemur
78	165	27.5	217685	2 AC149677	Bos tauru
79	159.2	26.5	340	6 AX898564	Sequence
80	159.2	26.5	340	6 BD034097	Sequence
81	154.6	25.7	2156	3 AK112519	Clona int
82	146.2	24.3	829	6 C0580257	Sequence
83	146.2	24.3	1821	3 AY095033	Drosophi1
84	126.4	21.0	252	6 AX898571	Sequence
85	126.4	21.0	252	6 BD034104	Sequence
86	116	19.3	4408	6 C0580256	Sequence
87	116	19.3	69208	2 AC020466	Drosophi1
88	116	19.3	175118	3 AC010842	Drosophi1
89	116	19.3	188272	3 AC005639	Drosophi1
90	116	19.3	295225	3 AE003461	Drosophi1
91	85.4	14.2	714	6 BD146713	Primer fo
92	85.4	14.2	714	6 AX866651	Sequence

93	85..4	14..2	1602	6	BD160707	Primer fo
94	85..4	14..2	1602	6	AX884081	Sequence
95	85..4	14..2	1602	6	AK024161	Homo sapi
96	60	10..0	60	6	CQ553517	Sequence
97	47..4	7..9	7095	6	AX700181	Sequence
98	47..4	7..9	7095	10	RRMAP1B5	X60370 R.norvegicu
99	47..4	7..9	225204	2	AC127668	Rattus no
100	47..4	7..9	229705	2	AC130178	Rattus no

ALIGNMENTS

RESULT 1
AF070671 1892 bp mRNA linear PRI 21-JUN-1999
DEFINITION Homo sapiens TNF-induced protein GSG-1 mRNA, complete cde.
ACCESSION AF070671
VERSION AF070671.1 GI:3978237
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE 1 (bases 1 to 1892)
Horrevorts,A.J., Fontijn,R.D., van Zonneveld,A.J., de Vries,C.J.,
ten Cate,J.W. and Pannekoek,H.
Vascular endothelial genes that are responsive to tumor necrosis
factor-alpha in vitro are expressed in atherosclerotic lesions,
including inhibitor of apoptosis protein-1, stannin, and two novel
genes
Blood 93 (10), 3418-3431 (1999)

JOURNAL
MEDLINE 99252096
PUBMED 10233894
REFERENCE 2 (bases 1 to 1892)
Horrevorts,A.J.G., Fontijn,R.D. and Pannekoek,H.
Direct Submission
Submitted (05-JUN-1998) Biochemistry, Academic Medical Center,
Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
JOURNAL location/Qualifiers
FEATURES
source
1..1892
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="endothelial"
/tissue_type="umbilical vein"
98..664
/codon_start=1
/product="TNF-induced protein GSG-1"
/protein_id="AAC83229.1"
/db_xref="GI:3978238"
/translation="MATDVFNKLNVAQAKILKMKVSKIATLLIDTSEVDEL
YRVTREYONKKEARKIKILIKYIKALILRNQFNDDELALMEKEKKVHQIAMS
VVFHQVDITFDNRNLSRLNCRMLHQIIOHLLAKSHKRVNVPFHPISCEFLAA
LNPNGNFKPHIQKLCDGINKMLDEENI"

CDS

ORIGIN
Query Match 97.9%; Score 588.4; DB 9; Length 1892;
Best Local Similarity 99.7%; Pred. No. 2.1e-136;
Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
1 CTTGACGCTCCGGCGCGCTGC-CGACTCTCCGATGCGCACAGATGCTTTAATCCA 59
2 CTTGACGCTCCGGCGCGCTGC-CGACTCTCCGATGCGCACAGATGCTTTAATCCA 122
3 CTTGACGCTCCGGCGCGCTGC-CGACTCTCCGATGCGCACAGATGCTTTAATCCA 122
4 CTTGACGCTCCGGCGCGCTGC-CGACTCTCCGATGCGCACAGATGCTTTAATCCA 119
5 CTTGACGCTCCGGCGCGCTGC-CGACTCTCCGATGCGCACAGATGCTTTAATCCA 182
6 CTTGACGCTCCGGCGCGCTGC-CGACTCTCCGATGCGCACAGATGCTTTAATCCA 179
7 CTTGACGCTCCGGCGCGCTGC-CGACTCTCCGATGCGCACAGATGCTTTAATCCA 242

QY	180	GGAGTACACCCCAAAACAGAGAGGAGGAGAGATCATCAAGAACCTCATCAGACAG	239
DB	243	GGAGTACACCCCAAAACAGAGAGGAGGAGAGATCATCAAGAACCTCATCAGACAG	302
QY	240	TCATCAAGCTGGCCATCTTTATAGGAATTAATCAATCAAGATGAGCTTACATTTGA	299
DB	303	TCATCAAGCTGGCCATCTTTATAGGAATTAATCAATCAAGATGAGCTTACATTTGA	362
QY	300	TGGAGAAATTTAAGAGAAAGTTCATCAGCTTCATACCGTGTGATTTCCATCAGG	359
DB	363	TGGAGAAATTTAAGAGAAAGTTCATCAGCTTCATACCGTGTGATTTCCATCAGG	422
QY	360	TGATTTATACCTTTGACCGAATGTGTATCAGGCTTTAATGAATGACAGAGATGC	419
DB	423	TGATTTATACCTTTGACCGAATGTGTATCAGGCTTTAATGAATGACAGAGATGC	482
QY	420	TGCACCAATATCATTCAGGCGCACTGACTGCCAAGTCAATGACGAGGTTAATATGCT	479
DB	483	TGCACCAATATCATTCAGGCGCACTGACTGCCAAGTCAATGACGAGGTTAATATGCT	542
QY	480	TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAATCCTTTGGAAATTTTA	539
DB	543	TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAATCCTTTGGAAATTTTA	602
QY	540	AACCCACTTACAAAACTATGTATGATGATCAACAAATGTGATGAGAGACATAT	599
DB	603	AACCCACTTACAAAACTATGTATGATGATCAACAAATGTGATGAGAGACATAT	662
QY	600	GA 601	
DB	663	GA 664	

RESULT 2	BD135356	1943 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD135356	110 human secretory proteins.			
DEFINITION	BD135356	110 human secretory proteins.			
ACCESSION	BD135356.1	GI:23230301			
VERSION	JP 2002508167-A/107.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1943)				
AUTHORS	Moore,P.A., Ruben,S.M., Carter,K.C., Shi,Y., Rosen,C.A., Soppet,D.R., Caou,H., Wei,Y.F., Florence,K., Duan,R.D., Florence,C., Greene,J.M., Feng,P., Ferris,A.M., Yu,G.L., Janat,F. and N.J.				
TITLE	110 human secretory proteins				
JOURNAL	Patent: JP 2002508167-A 107 19-MAR-2002;				
COMMENT	HUMAN GENOME SCIENCES INC				
	OS Homo sapiens (human)				
	PN JP 2002508167-A/107				
	PD 19-MAR-2002				
	PF 17-DEC-1998 JP 2000539040				
	PR 18-DEC-1997 US 60/070 923,18-DEC-1997 US 60/068 007 PR				
	18-DEC-1997 US 60/068 057,18-DEC-1997 US 60/068 006 PR				
	18-DEC-1997 US 60/068 008,18-DEC-1997 US 60/068 054 PR				
	18-DEC-1997 US 60/068 064,18-DEC-1997 US 60/068 053 PR				
	19-DEC-1997 US 60/068 169,19-DEC-1997 US 60/068 368 PR				
	19-DEC-1997 US 60/068 367,19-DEC-1997 US 60/068 369 PR				
	19-DEC-1997 US 60/068 365				
	PI PAUL A MOORE,STEVEN M RUBEN,KENNETH C CARTER,YANGU SHI,CRAIG				
	PI A ROSEN,				
	PI DANIEL R SOPPET,HARA CAOU,YING FEI WEI,KIMBERLY FLORENCE, PI				
	ROSAANNE D DUAN,				
	PI CHARLES FLORENCE,JOHN M GREENE,PING FENG,ANN M FERRIE,GUO PI				
	LIANG YU,				
	PI FORD JANAT,JIAN NI				
	PI C12N15/09,A61K38/00,A61K48/00,A61P9/00,A61P9/10,A61P15/00, PC				
	A61P25/00,				
	PC A61P25/02,A61P25/14,A61P25/16,A61P25/18,A61P25/22,A61P25/24,				

|||||
Db 645 AACCCACTTACAAAACCTATGTGATGTATCAACAAAATGTTGATGAGAGACATAT 704
Qy 600 GA 601
Db 705 GA 706
RESULT 4
AX877846 1921 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 12751 from Patent EP1074617.
DEFINITION AX877846
ACCESSION AX877846
VERSION AX877846.1 GI:40032582
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12751 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source 1..1921
Location/Qualifiers
1..1921
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
140..706
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89903.1"
/db_xref="GI:40032583"
/translation="MATDVPSNKLVLAQAKILGKMSKSIATTLIDPSSSEVLDEL
/translation="MATDVPSNKLVLAQAKILGKMSKSIATTLIDPSSSEVLDEL
VYTRREYONKKEAEKIKILIKVILKILVYNNQFNDELAEMKFKKXVHOLAMT
VVSFHOVDYTPDRNYSRLINCEKREMLHOLIQRLTKASHGRNNVFHDHPSCEFLAA
LYNPNRNPFRHIOKLCIDGINKLDEMENT"
ORIGIN
Query Match 97.6%; Score 586.8; DB 6; Length 1921;
Best Local Similarity 99.5%; Pred. No. 5,3e-136;
Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 CTTGACCGTCCCGCGCCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATCCA 59
Db 105 CTTGACCGTCCCGCGCCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATCCA 164
Qy 60 AAAACCTGGCCGCTTCAAGCAAAAAGATCTTGGTAAATGGTGCMAATTCATCG 119
Db 165 AAAACCTGGCCGCTTCAAGCAAAAAGATCTTGGTAAATGGTGCMAATTCATCG 224
Qy 120 CCACCACTTAATAGACGACACAAGTAGAGGTGCTGATGAGCTCTACAGAGTACCA 179
Db 225 CCACCACTTAATAGACGACACAAGTAGAGGTGCTGATGAGCTCTACAGAGTACCA 284
Qy 180 GGGAGTACACCCAAAACAAGAGAGGACAGAAAGTCAAGAACTTCATCAAGACG 239
Db 285 GGGAGTACACCCAAAACAAGAGAGGACAGAAAGTCAAGAACTTCATCAAGACG 344
Qy 240 TCATCAAGCGGCGCATCTTTATAGAAATPACATGTTAATCAAGATGAGCTAGCATTA 299
Db 345 TCATCAAGCGGCGCATCTTTATAGAAATPACATGTTAATCAAGATGAGCTAGCATTA 404
Qy 300 TGGAGAAATTTAAGAAAGTTCATCAGCTTGCTATGACCGTGTCAAGTTCCATCAG 359
Db 405 TGGAGAAATTTAAGAAAGTTCATCAGCTTGCTATGACCGTGTCAAGTTCCATCAG 464
Qy 360 TGGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGATGACAGAGATGC 419
Db 465 TGGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGATGACAGAGATGC 524

Qy 420 TGCACCAATCATTCAGCGCCACCTCACTGCCAAGTCACATGACGGGTAAATATGCT 479
Db 525 TGCACCAATCATTCAGCGCCACCTCACTGCCAAGTCACATGACGGGTAAATATGCT 584
Qy 480 TTGATCATTTTTACAGATTGTGAATTTTGGCTGCTGTATATCCTTTGGGAATTTTA 539
Db 585 TTGATCATTTTTACAGATTGTGAATTTTGGCTGCTGTATATCCTTTGGGAATTTTA 644
Qy 540 AACCCACTTACAAAACCTATGTGATGTATCAACAAAATGTTGATGAGAGACATAT 599
Db 645 AACCCACTTACAAAACCTATGTGATGTATCAACAAAATGTTGATGAGAGACATAT 704
Qy 600 GA 601
Db 705 GA 706
RESULT 5
AK001931 1921 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ11069, f1s, clone PLACE1004930, highly similar
DEFINITION to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK001931
VERSION AK001931.1 GI:7023502
KEYWORDS oligo cloning; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Makamatsu,A., Hayashi,K., Saito,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahata,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shitatori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuma,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Nishimura,K., Ishihashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hirooka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Houchuta,T., Kusano,Y., Kanehori,K., Takahashi-Fuji,A.,
Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Mueseshino,K., Yuzuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hiro,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Maehuo,Y., Yamasaki,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
TITLE Complete sequencing and characterization of 21,243 full-length
human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702039
REFERENCE
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y.,
Matsunawa,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Makamatsu,A., Nakamura,Y., Negahata,K., Maehuo,Y. and Sasaki,N.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 1921)
REFERENCE Isogai,T. and Otsuki,T.
TITLE Direct Submission

JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yama, Kisarazu, Chiba 252-0812, Japan (E-mail:genomics@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
 .source
 Location/Qualifiers
 1.1921
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1004930"
 /issue_type="placenta"
 /clone_lib="PLACE1"
 /note="cloning vector: pME185FL3"

ORIGIN
 Query Match 97.6%; Score 586.8; DB 9; Length 1921;
 Best Local Similarity 99.5%; Pred. No. 5.3e-136;
 Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 CTTGACGCTCCGGCGCGCTGCG-CGACTCTCCGATGGCCACAGATGCTTTAATTC 59
 105 CTTTCAGCGTCCCGCGCGCGCGCCCTCCGATGGCCACAGATGCTTTAATTC 164
 60 AAAACCTGGCGCTTCAAGGCAAAAGAGATCTTGGTAAATGGTCCAAATTCATCG 119
 165 AAAACCTGGCGCTTCAAGGCAAAAGAGATCTTGGTAAATGGTCCAAATTCATCG 224
 120 CCACACCTTAATAGACACACAGATGAGGTGCTGATGAGCTCTACAGATGACCA 179
 225 CCACACCTTAATAGACACAGATGAGGTGCTGATGAGCTCTACAGATGACCA 284
 180 GGGAGTACACCCCAAAAGAGAGAGGAGAGATCATCAAGAACTTCATCAAGACAG 239
 285 GGGAGTACACCCCAAAAGAGAGAGGAGAGATCATCAAGAACTTCATCAAGACAG 344
 240 TCATCAAGCTGGCTCTTATAGAAATATAGTTATCAAGATGAGCTGACATTCGA 239
 345 TCATCAAGCTGGCTCTTATAGAAATATAGTTATCAAGATGAGCTGACATTCGA 404
 300 TGGAGAAATTTAAGAAAGTTCATCAGCTGCTGATGACCGTGTCAAGTTCCATCAGG 359
 405 TGGAGAAATTTAAGAAAGTTCATCAGCTGCTGATGACCGTGTCAAGTTCCATCAGG 464
 360 TGGATTATACCTTTGACCGGAATGTGTATCAAGCTGTTAAATGAATGACAGAGATGC 419
 465 TGGATTATACCTTTGACCGGAATGTGTATCAAGCTGTTAAATGAATGACAGAGATGC 524
 420 TGGACCAATTCATTCAGCGCCACCTCACTGCCAAGTCACTGAGCGGTTAAATGTC 479
 525 TGGACCAATTCATTCAGCGCCACCTCACTGCCAAGTCACTGAGCGGTTAAATGTC 584
 480 TTGATCATTTTTCAGATTTGATTTTGGCTGCTGTTAATCCTTTGGGAATTTTA 539
 585 TTGATCATTTTTCAGATTTGATTTTGGCTGCTGTTAATCCTTTGGGAATTTTA 644
 540 AACCCCACTTAACTAATGTGATGATTCACAAATGTTGATGAAGAGACATAT 599
 645 AACCCCACTTAACTAATGTGATGATTCACAAATGTTGATGAAGAGACATAT 704
 600 GA 601
 705 GA 706

RESULT 6
 BD156785 LOCUS BD156785 1729 bp DNA linear PART 17-JAN-2003

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD156785
 VERSION BD156785.1 GI:27862543
 KEYWORDS JP 2002191363-A/11628.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1729)
 Ota,T., Isogai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 11628 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/11628
 PD 09-JUL-2002
 PP 28-JUL-2000 JP 200280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 (71)..(664).
 FT CDS
 1.1729
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 93.9%; Score 564.4; DB 6; Length 1729;
 Best Local Similarity 99.8%; Pred. No. 2.2e-130;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCACAAACCTGGCGCTTACAGGACCAAAAGAAATCTTGG 95
 102 TGGCCACAGATGCTTTAATTCACAAACCTGGCGCTTACAGGACCAAAAGAAATCTTGG 161
 96 GTAAATGCTGTCAAATTCATCGCCACCACTTAATAGACACACAAGTAGAGATGC 155
 162 GTAAATGCTGTCAAATTCATCGCCACCACTTAATAGACACACAAGTAGAGATGC 221
 156 TGGATGAGCTTACAGAGTGAACGAGGATGACACCAAAACAAAGAGAGGACAGAGA 215
 222 TGGATGAGCTTACAGAGTGAACGAGGATGACACCAAAACAAAGAGAGGACAGAGA 281
 216 TCATCAAGAACTTCATCAAGACAGCATCAAGTGGCCATCTTTAGGAATTAACAG 275
 282 TCATCAAGAACTTCATCAAGACAGCATCAAGTGGCCATCTTTAGGAATTAACAG 341
 276 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGATTCATCAGCTGCTA 335
 342 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGATTCATCAGCTGCTA 401
 336 TGAACGCTGCTGATTCATCAGTGAATTAATCACTTGAACCGGAATGTGTTATCCAGGC 395
 402 TGAACGCTGCTGATTCATCAGTGAATTAATCACTTGAACCGGAATGTGTTATCCAGGC 461
 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCACTTGAAGGCCACCTCACTGCCAAG 455
 462 TGTAAATGAATGACAGAGATGCTGACCAAAATCACTTGAAGGCCACCTCACTGCCAAG 521
 456 CACATGACGGGTTAATAATGCTTTGATCATTTTTCAGATTTGATGATTTTGGCTGCT 515
 522 CACATGACGGGTTAATAATGCTTTGATCATTTTTCAGATTTGATGATTTTGGCTGCT 581
 516 TGTATATCCTTTTGGGAAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAACA 575

Db	582	TGTAATATCCCTTTTGGGAATTTTAAACCCCACTTACAAAAACTATGTATGTATCAACA	641
Qy	576	AAATGTTGATGAAGAGAACTATGA	601
Db	642	AAATGTTGATGAAGAGAACTATGA	667
RESULT 7			
AX877673			
LOCUS	AX877673	1729 bp	DNA
DEFINITION	Sequence 12578 from Patent EP1074617.		linear
ACCESSION	AX877673		
VERSION	AX877673.1	GI:40032409	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Mumaiyotai; Metaxoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumalia; Eutheria; Primates; Carniaria; Homiidae; Homo.		
TITLE	Ort.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J., Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.		
JOURNAL	Primers for synthesizing full-length cDNA and their use		
FEATURES	Patent: EP 1074617-A 12578 07-FEB-2001; Research Association for Biotechnology (JP)		
source	location/Qualifiers		
	1..1729		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	71..667		
	/note="unassigned protein product"		
	/codon_start=1		
	/protein_id="CAE89829.1"		
	/db_xref="GI:40032410"		
	/translation="MHSEESKEVADVENSKNLAVQOKKILGKRVSKIATTLID		
	DTSEVDELRYVTRREYONKKEAKIIKNIKTVIKALIRNNQFODELATMEKFK		
	KKHOHLATVYSFHOVDYTPDRNVLSSLNECEHMLQIIRHLTAKSHGRVANNVDF		
	HFSDCEFLAALNPFGNPKPHLOKLCIDGIMGDEBENT"		
ORIGIN			
Query Match	93.9%; Score 564.4; DB 6; Length 1729;		
Best Local Similarity	99.8%; Pred. No. 2.2e-130;		
Matches 565; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Qy	36	TGGCCACGAGTCTTTAATCCAAAACCTGGCCGTCAGGCACAAAGAAGATCTTG	95
Db	102	TGGCCACGAGTCTTTAATCCAAAACCTGGCCGTCAGGCACAAAGAAGATCTTG	161
Qy	96	GTAAGAATGATGCCAATTCATCGCCACCACTTATAGACGACCAAGTAGTAGTGC	155
Db	162	GTAAGAATGATGCCAATTCATCGCCACCACTTATAGACGACCAAGTAGTAGTGC	221
Qy	156	TGATGACCTCTACAGAGTGACCCAGGAGTACACCCAAAACAGAGGACAGAGAGA	215
Db	222	TGATGACCTCTACAGAGTGACCCAGGAGTACACCCAAAACAGAGGACAGAGAGA	281
Qy	216	TCATCAAAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAGT	275
Db	282	TCATCAAAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAGT	341
Qy	276	TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAGAAATTCATCAGCTTGCTA	335
Db	342	TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAGAAATTCATCAGCTTGCTA	401
Qy	336	TGACCGTGTCAGTTTCATCAGTGTGATTTATCCTTTGACCGGATGTATTATCAGGC	395
Db	402	TGACCGTGTCAGTTTCATCAGTGTGATTTATCCTTTGACCGGATGTATTATCAGGC	461
Qy	396	TGTTAAATGAATGACAGAGATGCTGCACAAATCATTCAGGCGCACTGCCTGCAGT	455
Db	462	TGTTAAATGAATGACAGAGATGCTGCACAAATCATTCAGGCGCACTGCCTGCAGT	521

Oy	456	CACATGAGCGGGTAAATAAAGTCCTTTGATCATTTTTGAGATTGTGAATTTTGGCTGCCT	515
Dd	522	CACATGAGCGGGTAAATAAAGTCGTGTGATCATTTTTGAGATTGTGAATTTTGGCTGCCT	581
Oy	516	TGTAAATCCTTTTGGGAATTTTAACCCCACTTACA AAAACTATGTGATGTATCAACA	575
Dd	582	TGTAAATCCTTTTGGGAATTTTAAACC CCACTTACAAA AACTATGTGATGTATCAACA	641
Oy	576	AAATGTGGATGAAGAGACATATAGA	601
Dd	642	AAATGTGGATGAAGAGACATATAGA	667
RESULT 8			
AKO01850			
LOCUS			
DEFINITION	AKO01850	1729 bp mRNA linear PRI 30-JAN-2004	
ACCESSION	AKO01850	Homo sapiens cDNA FLJ10986 fis, clone PLACEH1001920, highly similar	
KEYWORDS	AKO01850	KO1850	
VERSIONS	AKO01850.1	GI:7023373	
RECORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
	1		
	Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,		
	Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,		
	Setino,M., Ohayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,		
	Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,		
	Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuna,M.,		
	Shitatori,A., Sudo,H., Hosobiri,T., Kaku,Y., Kodaira,H., Kondo,H.,		
	Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,		
	Kikkawa,E., Omura,Y., Abe,K., Kimihara,K., Kakesuta,N., Sato,K.,		
	Tanikawa,M., Yamazaki,M., Niimomiya,K., Ishibashi,T., Yamashta,H.,		
	Murakawa,K., Fujimori,K., Tanaai,H., Kimata,M., Matanabe,M.,		
	Hiraocha,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Matanabe,S.,		
	Yosida,M., Hotta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,		
	Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,		
	Takenuchi,K., Arita,M., Imose,N., Muraashino,K., Yuuki,H., Oshima,A.,		
	Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,		
	Shohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,		
	Terasima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,		
	Goto,Y., Shimizu,F., Wakebe,H., Hisigaki,H., Watanabe,T.,		
	Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,		
	Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,		
	Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,		
	Ozaki,K., Hitano,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,		
	Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,		
	Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsunaga,K.,		
	Nakaijima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Tohashi,T.,		
	Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,		
	Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,		
	Nagase,T., Nomura,N., Kikuchi,H., Masuno,Y., Yamashta,R.,		
	Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.		
	Complete sequencing and characterization of 21,243 full-length		
	human cDNAs		
TITLE			
JOURNAL			
PUBMED			
REFERENCE	Nat. Genet. 36 (1), 40-45 (2004)		
AUTHORS	14702039		
	2		
	Isogai,T., Oca,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,		
	Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,		
	Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,		
	Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,		
	Wakamatsu,A., Nakamura,Y., Nagahari,K., Maeno,Y. and Oshima,A.		
	NEO human cDNA sequencing project		
TITLE	Unpublished		
JOURNAL			
REFERENCE	3 (bases 1 to 1729)		
AUTHORS	Isogai,T. and Otsuki,T.		
TITLE	Direct Submission		
JOURNAL			
	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,		
	Genomics Laboratory/ 1532-3 Yana, Kisarazu, Chiba 229-0812, Japan		
	(E-mail:genoml@chit.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		

COMMENT

NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES

Location/Qualifiers
1..1729

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLAC1001920"
/tissue_type="placenta"
/clone_id="PLAC1"
/note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 1729;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTAAATTCACAAAACCTGGCCGTTCAAGCACAAGAAAGATCTTGG 95
102 TGGCCACAGATGCTTAAATTCACAAAACCTGGCCGTTCAAGCACAAGAAAGATCTTGG 161
96 GTAAATGATGCTTCAATTCATGCGCAACCTTAATAGACACACAAGTAGTAGGTGC 155
162 GTAAATGATGCTTCAATTCATGCGCAACCTTAATAGACACACAAGTAGTAGGTGC 221
156 TGGATGAGCTTACAGAGTACCAAGGAGTACACCCAAAACAAGAGAGGACAGAGAGA 215
222 TGATGAGCTTACAGAGTACCAAGGAGTACACCCAAAACAAGAGAGGACAGAGAGA 281
216 TCATCAGAACCTCATCAAGACAGTATCAAGCTGGCCATCTTATAGGAATATAGT 275
282 TCATCAGAACCTCATCAAGACAGTATCAAGCTGGCCATCTTATAGGAATATAGT 341
276 TTAATCAAGTGAAGTACATGATGAGAAATTTAAGAAAGATTCATCAGCTGCTCA 335
342 TTAATCAAGTGAAGTACATGATGAGAAATTTAAGAAAGATTCATCAGCTGCTCA 401
336 TGAACCTGCTCAGTTCATCAGAGTGAATATACCTTTGACCGGAATGTGTATCCAGGC 395
402 TGAACCTGCTCAGTTCATCAGAGTGAATATACCTTTGACCGGAATGTGTATCCAGGC 461
396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCACCTCAGTCCAGAT 455
462 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCACCTCAGTCCAGAT 521
456 CACATGAGACGGGTTAATATATGCTTTGATCATTTTTCAGATTTTGGCTGCTCT 515
522 CACATGAGACGGGTTAATATATGCTTTGATCATTTTTCAGATTTTGGCTGCTCT 581
516 TGTAAATGCTTTTGGAAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 575
582 TGTAAATGCTTTTGGAAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 641
576 AAATGTTGATGAAGAAACATATGA 601
642 AAATGTTGATGAAGAAACATATGA 667

RESULT 9
BC005352 1814 bp mRNA linear PRI 24-FEB-2004
LOCUS Homo sapiens tumor necrosis factor, alpha-induced protein 8, mRNA
DEFINITION (CDNA clone MGC:12451 IMAGE:3997650), complete cds.
ACCESSION BC005352
VERSION BC005352.1 GI:13529163
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 1814)
Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner K.H., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarnate P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bonfard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schultz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2. (bases 1 to 1814)
Straussberg R.
Direct Submission
Submitted (27-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomes Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (ILNLI)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxi.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

REMARK

COMMENT

FEATURES

source

gene

CDS

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: <http://image.llnl.gov>
Series: IRAL Plate: 16 Row: D Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657123.
Location/Qualifiers
1..1814
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12451 IMAGE:3997650"
/tissue_type="Bladder, carcinoma"
/clone_lib="NIH MGC_53"
/lab_host="PH105"
/note="Vector: pDNR-LIB"
1..1814
/gene="TNFAIP8"
/note="synonyms: G92-1, MDC-3.13, SCC-S2"
/db_xref="locusid:25816"
144..740
/gene="TNFAIP8"
/codon_start=1
/product="TNFAIP8 protein"
/protein_id="AAH05352.1"
/db_xref="GI:13529164"
/db_xref="locusid:25816"
/translation="MSEAESESKVAATDVFNKSNLAVQAKILGKMSKSTATTLID
DTSSVLDLRYVTRBYTQNKAEKIKLIKTYIKAILYRNNOFNDELALMEKF

```

misc_feature
      KKKVQLAMTVVSPHQVDYTFDRNVL SRLINCREMTHQI IORHLTAKSHGRVNVFD
      HSDCEFLAALYNPFGNFKPHLOKLC DGINKMLDEBNT
      180..737
      /gene="TMFAP18"
      /note="DUF758; Region: Domain of unknown function
      (DUF758). Family of eukaryotic proteins with unknown
      function, which are induced by tumour necrosis factor"
      /db_xref="CD:pfam05527"

ORIGIN
Query Match      93.9%; Score 564.4; DB 9; Length 1814;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      36  TGGCCACAGATGCTTTAATTCCTGCGCGCTGAGGACAAAAGAGATCTTG 95
      175  TGGCCACAGATGCTTTAATTCCTGCGCGCTGAGGACAAAAGAGATCTTG 234
QY      96  GTAAAATGGTGTCCAATTCATGCGCACAACCTTAATAGACACAGAGTAGAGTGC 155
      235  GTAAAATGGTGTCCAATTCATGCGCACAACCTTAATAGACACAGAGTAGAGTGC 294
QY      156  TGGATGAGCTTACAGAGTACCGAGGAGTACACCCAAAACAAGAGGACAGAGA 215
      295  TGGACGAGCTTACAGAGTACCGAGGAGTACACCCAAAACAAGAGGAGGAGAGA 354
QY      216  TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATTCTTATAGGAATATCAGT 275
      355  TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATTCTTATAGGAATATCAGT 414
QY      276  TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCAACGCTTGCTA 335
      415  TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCAACGCTTGCTA 474
QY      336  TGAACCTGTGCTTCCATCAAGTGGATTAATCCTTTGACCGGATGTGTTATCCAGG 395
      475  TGAACCTGTGCTTCCATCAAGTGGATTAATCCTTTGACCGGATGTGTTATCCAGG 534
QY      396  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCACCTCATGCCAAGT 455
      535  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCACCTCATGCCAAGT 594
QY      456  CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTGAAATTTTGGCTGCT 515
      595  CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTGAAATTTTGGCTGCT 654
QY      516  TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 575
      655  TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 714
QY      576  AAATGTTGATGAAGAAACATATGA 601
      715  AAATGTTGATGAAGAAACATATGA 740
      Db

RESULT 10
AF099935      1915 bp      mRNA      linear      PRI 12-NOV-1998
LOCUS      Homo sapiens MDC-3.13 isoform 2 mRNA, complete cds.
DEFINITION      AF099935
ACCESSION      AF099935
VERSION      AF099935.1 GI:3860092
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Niefeld,W. and Meyerhans,A.F.
TITLE      Identification of cellular factors involved in the differentiation
of dendritic cells
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 1915)
AUTHORS      Niefeld,W. and Meyerhans,A.F.

```

```

TITLE      Direct Submission
JOURNAL      Submitted (13-OCT-1998) Max Planck Institute for Molecular
Genetics, Immesstrasse 73, Berlin 14195, Germany
FEATURES
      source
      1..1915
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /cell_type="adherent monococytes treated with GM-CSF and
      IL-4"
      85..681
      /note="MDC-3.13/2"
      /codon_start=1
      /product="MDC-3.13 isoform 2"
      /protein_id="AAC72975.1"
      /db_xref="GI:3860093"
      /translation="HSDABESKEVATVPNSKILAVQAKKILGKMSKSLATLID
      DTSSVLDELKRVREYQNKKEAKKIKLITVTKIALLIRNNQFODELALMEKF
      KKKVQLAMTVVSPHQVDYTFDRNVL SRLINCREMTHQI IORHLTAKSHGRVNVFD
      HSDCEFLAALYNPFGNFKPHLOKLC DGINKMLDEBNT"

ORIGIN
Query Match      93.9%; Score 564.4; DB 9; Length 1915;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      36  TGGCCACAGATGCTTTAATTCCTGCGCGCTGAGGACAAAAGAGATCTTG 95
      116  TGGCCACAGATGCTTTAATTCCTGCGCGCTGAGGACAAAAGAGATCTTG 175
QY      96  GTAAAATGGTGTCCAATTCATGCGCACAACCTTAATAGACGACAAGTAGTAGTGC 155
      176  GTAAAATGGTGTCCAATTCATGCGCACAACCTTAATAGACGACAAGTAGTAGTGC 235
QY      156  TGGATGAGCTTACAGAGTACCGAGGAGTACACCCAAAACAAGAGGACAGAGA 215
      236  TGGATGAGCTTACAGAGTACCGAGGAGTACACCCAAAACAAGAGGAGGAGAGA 295
QY      216  TCATCAAGAACCTCATCAAGACAGTCAACAGCTGGCCATTCTTATAGGAATATCAGT 275
      296  TCATCAAGAACCTCATCAAGACAGTCAACAGCTGGCCATTCTTATAGGAATATCAGT 355
QY      276  TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCAACGCTTGCTA 335
      356  TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCAACGCTTGCTA 415
QY      336  TGAACCTGTGCTTCCATCAAGTGGATTAATCCTTTGACCGGATGTGTTATCCAGG 395
      416  TGAACCTGTGCTTCCATCAAGTGGATTAATCCTTTGACCGGATGTGTTATCCAGG 475
QY      396  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCACCTCATGCCAAGT 455
      476  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCACCTCATGCCAAGT 535
QY      456  CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTGAAATTTTGGCTGCT 515
      536  CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTGAAATTTTGGCTGCT 595
QY      516  TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 575
      596  TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 655
QY      576  AAATGTTGATGAAGAAACATATGA 601
      656  AAATGTTGATGAAGAAACATATGA 681
      Db

RESULT 11
C0726075      1988 bp      DNA      linear      PAT 03-FEB-2004
LOCUS      Sequence 12009 from Patent WO02068579.
DEFINITION      C0726075
ACCESSION      C0726075
VERSION      C0726075.1 GI:42288070

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanecons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 12009 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers
1. 1988
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 93.9%; Score 564.4; DB 6; Length 1988;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTAATTCACAAAACCTGGCCGCTTCAGGACCAAAAGAAGATCTTG 95
205 TGGCCACAGATGCTTAATTCACAAAACCTGGCCGCTTCAGGACCAAAAGAAGATCTTG 264
96 GTAAATGATGCTCAATTCATGCGCCACCACTTAATAGACGACAAAGTAGAGTG 155
265 GTAAATGATGCTCAATTCATGCGCCACCACTTAATAGACGACAAAGTAGAGTG 324
156 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGAGGAGAGA 215
325 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGAGGAGAGA 384
216 TCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGAAATATCAGT 275
385 TCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGAAATATCAGT 444
276 TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 335
445 TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 504
336 TGAACCTGCTCACTTCATCAAGTGTGATTAATCTTTGACCGGAATGTGTTATCCAGG 395
505 TGAACCTGCTCACTTCATCAAGTGTGATTAATCTTTGACCGGAATGTGTTATCCAGG 564
396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTCAGCGCCCACTCACTGCCAAGT 455
565 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTCAGCGCCCACTCACTGCCAAGT 624
456 CACATGAGAGGGTTAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCC 515
625 CACATGAGAGGGTTAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCC 684
516 TGTAAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATATCAACA 575
685 TGTAAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATATCAACA 744
576 AAATGTTGATGAAGAACAATATGA 601
745 AAATGTTGATGAAGAACAATATGA 770

RESULT 12
AF099936 2003 bp mRNA linear PRI 12-NOV-1998
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens MDC-3.13 isoform 1 mRNA, complete cds.
AF099936.1 GI:3860094
Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Nietfeld, W. and Meyerhans, A.F.
Identification of cellular factors involved in the differentiation
of dendritic cells
Unpublished
2. (bases 1 to 2003)
Nietfeld, W. and Meyerhans, A.F.
Direct Submision
Submitted (13-OCT-1998) Max Planck Institute for Molecular
Genetics, Ihmestrasse 73, Berlin 14195, Germany
Location/Qualifiers
1. 2003
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="adherent monocytes treated with GM-CSF and
IL-4"
198. 770
/note="MDC-3.13/1"
/product="MDC-3.13 isoform 1"
/codon_start=1
/protein_id="AA072976.1"
/db_xref="GI:3860095"
/translation="MAVATDVNSKNLAVQAKILGKVSIAITTLIDTNSSEYLD
ELRYRREYDTNKKKEAEKIKILIKVILALILYRNQNOBELAMEFKKHVOLA
MIVVSPHOVDYDFDRVLSRLINCEGHEHQLIQRHLTKSHQRVNVVDFSDCSFL
AALVNPFGFKHLOKLDGINKMDENI"

Query Match 93.9%; Score 564.4; DB 9; Length 2003;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTAATTCACAAAACCTGGCCGCTTCAGGACCAAAAGAAGATCTTG 95
205 TGGCCACAGATGCTTAATTCACAAAACCTGGCCGCTTCAGGACCAAAAGAAGATCTTG 264
96 GTAAATGATGCTCAATTCATGCGCCACCACTTAATAGACGACAAAGTAGAGTG 155
265 GTAAATGATGCTCAATTCATGCGCCACCACTTAATAGACGACAAAGTAGAGTG 324
156 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGAGGAGAGA 215
325 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGAGGAGAGA 384
216 TCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGAAATATCAGT 275
385 TCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGAAATATCAGT 444
276 TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 335
445 TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 504
336 TGAACCTGCTCACTTCATCAAGTGTGATTAATCTTTGACCGGAATGTGTTATCCAGG 395
505 TGAACCTGCTCACTTCATCAAGTGTGATTAATCTTTGACCGGAATGTGTTATCCAGG 564
505 TGAACCTGCTCACTTCATCAAGTGTGATTAATCTTTGACCGGAATGTGTTATCCAGG 564
396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTCAGCGCCCACTCACTGCCAAGT 455
565 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTCAGCGCCCACTCACTGCCAAGT 624
456 CACATGAGAGGGTTAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCC 515
625 CACATGAGAGGGTTAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCC 684
516 TGTAAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATATCAACA 575
685 TGTAAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATATCAACA 744
576 AAATGTTGATGAAGAACAATATGA 601

CDs

Db 745 AATGTTGGATGAAGAACATATGA 770

RESULT 13
BC007014

LOCUS

DEFINITION BC007014 2081 bp mRNA linear PRI 24-FEB-2004
Homo sapiens tumor necrosis factor, alpha-induced protein 8, mRNA
(cDNA clone MGC:12346 IMAGE:3930240), complete cds.

ACCESSION BC007014

VERSION BC007014.1 GI:13937825

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2081)
Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carrincci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skelske, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE 2 (bases 1 to 2081)

JOURNAL Strausberg, R.

PUBMED Direct Submission

REFERENCE Submitted (30-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

AUTHORS NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
source
1..2081
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12346 IMAGE:3930240"
/issue_type="Brain, Primitive neuroectodermal"
/clone_lib="NIH_MGC_56"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

gene 1..2081
/gene="TNFAIP8"
/note="synonyms: GG2-1, MDC-3.13, SCC-S2"
/db_xref="LocusID:25816"
237..833
/gene="TNFAIP8"
/codon_start=1
/product="TNFAIP8 protein"
/protein_id="AAH07014.1"
/db_xref="GI:13937825"
/db_xref="LocusID:25816"
/translation="MSEASESKVATDVNSKNLAVOAKKIKGWVKSIAITLLID
DTSEVDELIVTFREYQNKKEAKIKNLIKVTKAILIRNNQFODEIALMEKF
KKKHQALMTVSPHQVDYTFDPAVLSRLNECREMHLQIIQRHLTAKSHGVNVFD
HRSDEFLAALYNPRGNKPKPLQKLCDSDINMLDEEN"
273..830
misc_feature
/gene="TNFAIP8"
/note="DUF758; Region: Domain of unknown function (DUF758). Family of eukaryotic proteins with unknown function, which are induced by tumour necrosis factor"
/db_xref="CDD:pfam05527"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 2081;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCCCTTCAGGCACAAAAGAGATCTTGG 95
|||||
268 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCCCTTCAGGCACAAAAGAGATCTTGG 327
|||||
96 GTAAATGGTGTCCAAATTCATCGCCACCACTTAATGACGACACAAGTAGTAGGTGC 155
|||||
328 GTAAATGGTGTCCAAATTCATCGCCACCACTTAATGACGACACAAGTAGTAGGTGC 387
|||||
156 TGGATGAGCTTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAG 215
|||||
388 TGGATGAGCTTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAG 447
|||||
216 TCATCAAGAACTCTCAATCAAGATCAATCAAGTCTTCTTTATAGGAATATCACT 275
|||||
448 TCATCAAGAACTCTCAATCAAGATCAATCAAGTCTTCTTTATAGGAATATCACT 507
|||||
276 TTAATCAAGATGAGTACATGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 335
|||||
508 TTAATCAAGATGAGTACATGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 567
|||||
336 TGACCGTGTGATGATTCATCAAGTGTGATTTATACCTTTGACCGAATGTTATCCAGGC 395
|||||
568 TGACCGTGTGATGATTCATCAAGTGTGATTTATACCTTTGACCGAATGTTATCCAGGC 627
|||||
396 TGTTAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCACTCATCGCAAGT 455
|||||
628 TGTTAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCACTCATCGCAAGT 687
|||||
456 CACATGAGACGGGTATATATGCTTGTGATCATTTTTCAGATTTGGAATTTTGGCTGCT 515
|||||
688 CACATGAGACGGGTATATATGCTTGTGATCATTTTTCAGATTTTGGCTGCT 747
|||||
516 TGTATATATCTTTTGGGAATTTTAAACCCCACTTCAAAAACATATGATGATCAACA 575
|||||
748 TGTATATATCTTTTGGGAATTTTAAACCCCACTTCAAAAACATATGATGATCAACA 807
|||||
576 AATGTTGGATGAAGAACATATGA 601
|||||
808 AATGTTGGATGAAGAACATATGA 833
|||||

RESULT 14
AK097284 2502 bp mRNA linear PRI 30-JAN-2004
LOCUS AK097284

DEFINITION Homo sapiens cDNA FLJ39965 f1s, clone SPLEN2027157, highly similar to Homo sapiens MDC-3.13 isoform 2 mRNA.

ACCESSION AK097284
VERSION AK097284.1 GI:21756982
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Makatsugu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Mahta, H., Sekine, M., Okeyashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahori, K., Murakami, K., Yasuda, T., Iwano, T., Watanabe, M., Shiratori, A., Sudo, H., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, S., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Nimomiya, K., Ishibashi, T., Yamauchi, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirao, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoda, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi, F., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A., Sasaki, N., Aotsuma, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohara, N., Sano, S., Moriya, S., Momiya, H., Sato, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumei, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Oaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Naguchi, S., Itoh, T., Shigeta, K., Senda, T., Matsunura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL 14702039
PUBMED
REFERENCE
AUTHORS

TITLE 2
JOURNAL Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hirao, S., Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotta, T., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T., Kimura, K., Matsumoto, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Watanabe, M., Takahashi, F., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahori, K., Masuno, Y., Nagai, K., and Isogai, T. NEDO human cDNA sequencing project Unpublished

TITLE 3 (bases 1 to 2502)
JOURNAL Isogai, T. and Yamamoto, J.
REFERENCE Direct Submission
AUTHORS Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@nri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'- end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.

FEATURES
source location/Qualifiers
1. 2502
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SPLEN2027157"
/tissue_type="spleen"

ORIGIN
Query Match 93.9%; Score 564.4; DB 9; Length 2502;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/note="cloning vector: pME18SFL3"

36 TGGCCACAGATGCTTTAATTCACAAACCTGCGCTTCAGGACACAAAGACTCTGG 95
720 TGGCCACAGATGCTTTAATTCACAAACCTGCGCTTCAGGACACAAAGACTCTGG 779
96 GTAAATGCTGCTCAATTCATGCGACCACTTAATGACGACAAAGAGAGAGTGC 155
780 GTAAATGCTGCTCAATTCATGCGACCACTTAATGACGACAAAGAGAGAGTGC 839
156 TGGATGAGCTCTACAGATGACGAGGAGAGACCCAAACAAAGAGAGAGAGAGA 215
840 TGGATGAGCTCTACAGATGACGAGGAGAGACCCAAACAAAGAGAGAGAGAGA 899
216 TCATCAAGAACTCATCAAGACGATCATCAAGCTGCGCATCTTTATAGGAATATCAGT 275
900 TCATCAAGAACTCATCAAGACGATCATCAAGCTGCGCATCTTTATAGGAATATCAGT 959
276 TTAATCAAGTATGCTATGATGATGAGAAATTTAAGAAAGATTCATCACTTCTCA 335
960 TTAATCAAGTATGCTATGATGATGAGAAATTTAAGAAAGATTCATCACTTCTCA 1019
336 TGACCGTGGTCACTTCATCAAGTATGATGATGATGATGATGATGATGATGATGAT 395
1020 TGACCGTGGTCACTTCATCAAGTATGATGATGATGATGATGATGATGATGATGAT 1079
396 TGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 455
1080 TGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
456 CACATGAGCGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 515
1140 CACATGAGCGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
516 TGTATATCTTTTGGAAATTTAAGCCCACTTACAAAGATGATGATGATGATGATGAT 575
1200 TGTATATCTTTTGGAAATTTAAGCCCACTTACAAAGATGATGATGATGATGATGAT 1259
QY 576 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
Db 1260 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1285

RESULT 15
AC035144
LOCUS 112626 bp DNA linear PRI 27-APR-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2220M12, complete sequence.
ACCESSION AC035144
VERSION AC035144.3 GI:13811899
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 112626)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 112626)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 112626)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) DOE Joint Genome Institute, 2800 Mitchell

COMMENT

Drive, Walnut Creek, CA 94598, USA
On Apr 27, 2001 this sequence version replaced gi:7712094.
Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.2.
STS Content:

SHGC-7119 G14211.

FEATURES

source

Location/Qualifiers
1..112626
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2220M12"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 112626;
Best Local Similarity 99.8%; Pred. No. 2.4e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAAATTCGCAAAACCTGGCCGTTCCAGGCACAAAGAGATCTTGG 95
65049 TGGCCACAGATGCTTTAAATTCGCAAAACCTGGCCGTTCCAGGCACAAAGAGATCTTGG 65108
96 GTPAAATGGTTCGCAATTCATGCGCACCCCTTAATAGACGACACAGATGAGGTGC 155
65109 GTPAAATGGTTCGCAATTCATGCGCACCCCTTAATAGACGACACAGATGAGGTGC 65168
156 TGGATGAGCTCTACAGATGACAGGAGATGACCCCAAAACAGAGAGAGAGAGAGA 215
65169 TGGATGAGCTCTACAGATGACAGGAGATGACCCCAAAACAGAGAGAGAGAGAGA 65228
216 TCATCAAGAACTTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCACT 275
65229 TCATCAAGAACTTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCACT 65288
276 TTAATCAAGATGAGCTACATGATGAGAAATTTAAGAAAGAGATTCATCACTGCTGTA 335
65289 TTAATCAAGATGAGCTACATGATGAGAAATTTAAGAAAGAGATTCATCACTGCTGTA 65348
336 TGAACGTGGTCACTTCATCAGATGATATATCTTTGACCGGAATGTTATCCAGGC 395
65349 TGAACGTGGTCACTTCATCAGATGATATATCTTTGACCGGAATGTTATCCAGGC 65408
396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCCACTCAGTCCAGAT 455
65409 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCCACTCAGTCCAGAT 65468
456 CACATGAGAGGGGTTAATATGCTTGTGATATTTTCAAGTTGTGAATTTTGGTGGCT 515
65469 CACATGAGAGGGGTTAATATGCTTGTGATATTTTCAAGTTGTGAATTTTGGTGGCT 65528
516 TGTATATCTCTTTGGGAATTTTAAACCCCACTTAACAAAATCATGATGATATCAAC 575
65529 TGTATATCTCTTTGGGAATTTTAAACCCCACTTAACAAAATCATGATGATATCAAC 65588
576 AAATGTTGATGAAGAAACATATGA 601
65589 AAATGTTGATGAAGAAACATATGA 65614

RESULT 16
AC026795/c 156277 bp DNA linear PRI 21-SEP-2001
LOCUS AC026795 Homo sapiens chromosome 5 clone CTD-2287L22, complete sequence.
DEFINITION AC026795
AC026795
AC026795 GI:15718543
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished

REFERENCE

2 (bases 1 to 156277)
DOE Joint Genome Institute.
Direct Submission

REFERENCE

3 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE

4 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (31-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

5 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

6 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

7 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

8 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

9 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

10 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

11 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

12 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

13 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

14 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

15 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

16 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

17 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

18 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

19 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

20 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

21 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

22 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

23 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

24 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

25 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

26 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

27 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

28 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

29 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

30 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

31 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA


```

Db      127813 CACATGACGCGTTAATATGTGTGATCATTTTCAGATTGATGATTTTGGCTCCT 127754
QY      516 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACCTATGTGATGCTATCACA 575
Db      127753 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACCTATGTGATGCTATCACA 127694
QY      576 AATGTGTGATGAAGACATATGA 601
Db      127693 AATGTGTGATGAAGACATATGA 127668

RESULT 17
AC138612/c 158057 bp DNA linear PRI 30-MAY-2003
LOCUS      AC138612
DEFINITION Homo sapiens BAC clone CTD-2293E22 from 2, complete sequence.
ACCESSION  AC138612
VERSION     AC138612.1 GI:27597054
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 158057)
            Sulston, J.E. and Wilson, R.
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
JOURNAL    99063792
MEDLINE    9847074
PUBMED     9847074
REFERENCE  2 (bases 1 to 158057)
            Tomlinson, C. and Bielicki, L.
            The sequence of Homo sapiens BAC clone CTD-2293E22
            Unpublished (2001)
JOURNAL    3 (bases 1 to 158057)
            Waterston, R.H.
            Direct Submission
            Submitted (11-JAN-2003) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
JOURNAL    4 (bases 1 to 158057)
            Wilson, R.
            Direct Submission
            Submitted (30-MAY-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL    ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapliens@wustl.wustl.edu
            Summary Statistics
            Center project name: H_MS2293E22
            -----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
Clone CTD-2293E22 is from a release of the human BAC library CTD.

```

```

FEATURES
Source
    The library contains cloned DNA from human sperm. See: Shizuya et
    al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al.,
    Genomics 34:213-8 (1996). The clone is available from Research
    Genetics, Inc. (http://www.resgen.com).
    VECTOR: pBelOBAC11
    Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1. 158057
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2"
   /clone="CTD-2293E22"
   /clone_1b="CTD"
4. 83
   /rpt_family="MIR"
repeat_region 300..719
   /rpt_family="ERV1"
repeat_region 1321..1459
   /rpt_family="L1"
repeat_region 1518..1547
   /rpt_family="AT_rich"
repeat_region 1964..1996
   /rpt_family="(TA)n"
repeat_region 2532..2575
   /rpt_family="A-rich"
repeat_region 2639..2841
   /rpt_family="L2"
repeat_region 2859..2886
   /rpt_family="AT_rich"
repeat_region 3179..3201
   /rpt_family="AT_rich"
repeat_region 4351..4541
   /rpt_family="AT_rich"
repeat_region 5128..5766
   /rpt_family="MIR"
repeat_region 5767..6328
   /rpt_family="L1"
repeat_region 6329..7499
   /rpt_family="L1"
repeat_region 7503..7872
   /rpt_family="L1"
repeat_region 7920..8106
   /rpt_family="MIR"
repeat_region 8621..8849
   /rpt_family="ERV1"
repeat_region 9113..9415
   /rpt_family="Alu"
repeat_region 12251..12442
   /rpt_family="L1"
repeat_region 14323..14361
   /rpt_family="AT_rich"
repeat_region 14630..14899
   /rpt_family="Alu"
repeat_region 15273..15541
   /rpt_family="L1"
repeat_region 15534..17744
   /rpt_family="L1"
repeat_region 18238..18436
   /rpt_family="MER2_type"
repeat_region 18725..18973
   /rpt_family="MER2_type"
repeat_region 19348..19664
   /rpt_family="L2"
repeat_region 19666..19723
   /rpt_family="(TAAG)n"
repeat_region 20144..20192
   /rpt_family="polypurine"
repeat_region 20419..20468

```

36 TGGCCACAGATGTCCTTAATTCCAAAAACCTGGCCGTTCAGGCACAAGAAGATCTTGG 95

Db	135467	TTGGCCACGATGTCCTTTATTTCCAAAACCTGGCCGTTCCAGGCACAAAAGAAATCTTGG	135408
QY	96	GTAAATATGATGTCCTCAATTCATTCGCCACCACTTATATAGACGACCAAGTAGTAGTGTC	155
Db	135407	GTAAATATGATGTCCTCAATTCATTCGCCACCACTTATATAGACGACCAAGTAGTAGTGTC	135348
QY	156	TGATAGAGCTCTACAGAGTGACACGAGGAGTACACCCAAAACAAGAAAGGCGAGAGAA	215
Db	135347	TGATAGAGCTCTACAGAGTGACACGAGGAGTACACCCAAAACAAGAAAGGCGAGAGAA	135288
QY	216	TCATCAAGACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAGT	275
Db	135287	TCATCAAGACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAGT	135228
QY	276	TTAATCAAGATGAGCTAGCTATGATGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGA	335
Db	135227	TTAATCAAGATGAGCTAGCTATGATGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGA	135168
QY	336	TGACCTGGTCAGATTTCCATCAGCTGGATTTATCCTTTGACCGGAATGTGTTATCCAGGC	395
Db	135167	TGACCTGGTCAGATTTCCATCAGCTGGATTTATCCTTTGACCGGAATGTGTTATCCAGGC	135108
QY	396	TGTTAAATGAATGCAGAGAGATGCTGCACCAATCATTGACGGCCCACTCAGTCCCAAGT	455
Db	135107	TGTTAAATGAATGCAGAGAGATGCTGCACCAATCATTGACGGCCCACTCAGTCCCAAGT	135048
QY	456	CACATGACGCGGTTAATATATGCTTTGATCATTTTTCAGATTTGTGAATTTTGGCTGCT	515
Db	135047	CACATGACGCGGTTAATATATGCTTTGATCATTTTTCAGATTTGTGAATTTTGGCTGCT	134988
QY	516	TGTTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAATCTATGTGATCAACA	575
Db	134987	TGTTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAATCTATGTGATCAACA	134928
QY	576	AAATGTTGATGCAAGAAACATATGA 601	
Db	134927	AAATGTTGATGCAAGAAACATATGA 134902	
RESULT 18			
LOCUS	CR457137		
DEFINITION	CR457137 Homo sapiens full open reading frame cDNA clone RZPD0834H127D for	573 bp mRNA linear	PR1 03-JUN-2004
KEYWORDS	gene TGFAP8, tumor necrosis factor, alpha-induced protein 8;		
ACCESSION	CR457137		
VERSION	CR457137.1	GI:48146390	
KEYWORDS	Full ORF shuttle clone, Gateway(TM), complete cds.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
TITLE	1 (bases 1 to 573)		
JOURNAL	Ebert, U., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.		
REFERENCE	Cloning of human full open reading frames in Gateway(TM) system		
AUTHORS	entry vector (pDONR201)		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 573)		
REFERENCE	Ebert, U., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.		
AUTHORS	Direct Submission		
TITLE	Submitted (03-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer		
JOURNAL	Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,		
COMMENT	Germany		
	RZPD, RZPD0834H127D, ORFNO. 1910		
	www.rzpd.de/cgi-bin/products/cl.cgi?cloneId=RZPD0834H127D RZPDLIB,		
	Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.		
	834		
	www.rzpd.de/cgi-bin/products/showlib.pl.cgi?response?libNo=834		
	Contact: Inge Ariart		
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,		
	Heidenbergweg 6, D-14059 Berlin, Germany		

Tel: +49 30 32639 100
Fax: +49 30 32639 111

www.rzpd.de

This clone is available from RZPD;
contact RZPD (customer.service@rzpd.de) for further information.
This CDS clone is a part of a collection of human full length
expression clones generated by RZPD.

This CDS has been cloned incl. stopcodon.

The CDS has been inserted into pDONR201 via a BP Clonase(TM)
reaction. Additional sequence has been added in front of the start
codon: att. .AAAAA GCA GGC (ATG).

The last base of the last coding triplet has been changed to T,
which might lead to an amino acid change at the C terminus of the
polypeptide.

The stop codon has been set to TAA followed by

TTAACTCAGCTTCTT. att. Compared to the reference sequence NM_014350
we did not find any amino acid exchanges.

Clone distribution: http://www.rzpd.de/products/orfclones/.

FEATURES

source

```
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPD834H127D"
/clone_1b="Human Full ORF Clones Gateway(TM) - RZPD"
/lab_hosts="DH10B"
/note="Vector: pDONR201, Site_1: attP1, Site_2: attP2"
1..573
/gene="TNFAIP8"
1..573
/gene="TNFAIP8"
/codon_start=1
/protein_id="CAG3418.1"
/db_xref="GI:48146391"
/translation="MAVATDVFNASKNLAVOAKIIGKRWYSKIATLLIDTSSEVLD
ELYRTRVYONKKEAEKIIKNLKTIVITLALYRNPNODESLAMERKPKYVHOLA
MYVSEPHOVDYTRDRAVLSRLNIECBEMHOLIQRHUTAKSHGRVANNVDFHSDECEFL
AALYNPGRNFKPLQKLCGGINMLDEBN1"
```

ORIGIN

Query Match 93.4%; Score 561.2; DB 9; Length 573;
Best Local Similarity 99.5%; Pred. No. 1.3e-129;
Matches 563; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
36 TGGCCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTTCAGGACAAAAGAGATCTTGG 95
8 TGGCCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTTCAGGACAAAAGAGATCTTGG 67
96 GTTAAATGGTGTCCAATTCATGCGCACCACTTAATAGACGACAAAGTAGAGGTGC 155
68 GTTAAATGGTGTCCAATTCATGCGCACCACTTAATAGACGACAAAGTAGAGGTGC 127
156 TGGATGAGCTCTACAGAGTGACAGGAGAGTACACCAAAAAGAGAGAGGAGAGAGA 215
128 TGGATGAGCTCTACAGAGTGACAGGAGAGTACACCAAAAAGAGAGAGGAGAGAGA 187
216 TCATCAAGAACTTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGGAATATCA 275
188 TCATCAAGAACTTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGGAATATCA 247
276 TTTATCAAGTAGTGCATTCATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 335
248 TTTATCAAGTAGTGCATTCATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 307
336 TGACCGTGTGCTTCATCAGTGAATATATACCTTTGACCGGAATGTTTATCCAGGC 395
308 TGACCGTGTGCTTCATCAGTGAATATATACCTTTGACCGGAATGTTTATCCAGGC 367
396 TGTAAATGAGATGACAGAGAGTCTGCACCAAAATCAATTCAGCGCCACCTTCACCAAGT 455
368 TGTAAATGAGATGACAGAGAGTCTGCACCAAAATCAATTCAGCGCCACCTTCACCAAGT 427
456 CACATGAGCGGTTAATATGCTTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
```

428 CACATGAGCGGTTAATATGCTTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 487

516 TGTATATCTTTTGGAAATTTTAAACCCCACTTACAAAACATATGATGATGATCAACA 575

488 TGTATATCTTTTGGAAATTTTAAACCCCACTTACAAAACATATGATGATGATCAACA 547

576 AATGTTGATGAGAGAAACATATGA 601

548 AATGTTGATGAGAGAAACATTTAA 573

RESULT 19

BD149395

LOCUS BD149395 816 bp DNA linear PART 17-JAN-2003

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD149395 GI:27855153

VERSION BD149395.1 GI:27855153

KEYWORDS JP 2002191363-A/4238.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Oca.T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A 4238 09-JUL-2002;

HELEX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/4238

PD 09-JUL-2002

PF 28-JUL-2000 JP 200280990

PI TOSHIO OYA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU

PI SATTO

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KENICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/

10, C12P21/02,C12O1/68//C12P21/08,G06F1/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH key

Location/Qualifiers

FT source 1..816

Location/Qualifiers

1..816

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 91.1%; Score 547.6; DB 6; Length 816;
Best Local Similarity 98.9%; Pred. No. 3.4e-126;
Matches 562; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

```
36 TGGCCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTTCAGGACAAAAGAGATCTTGG 95
102 TGGCCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTTCAGGACAAAAGAGATCTTGG 161
96 GTTAAATGGTGTCCAATTCATGCGCACCACTTAATAGACGACAAAGTAGAGGTGC 155
162 GTTAAATGGTGTCCAATTCATGCGCACCACTTAATAGACGACAAAGTAGAGGTGC 221
156 TGGATGAGCTCTACAGAGTGACAGGAGAGTACACCAAAAAGAGAGAGGAGAGAGA 215
222 TGGATGAGCTCTACAGAGTGACAGGAGAGTACACCAAAAAGAGAGAGGAGAGAGA 281
216 TCATCAAGAACTTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGGAATATCA 275
282 TCATCAAGAACTTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGGAATATCA 341
276 TTTATCAAGATGAGCTGATGATGAGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 335
```

Db 342 TTAATCAATGATGCTAGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 401
 Qy 336 TGAACGCTGCTAGCTTCCATCAGGTGATTATACCTTTGACCGGAATGTGTATCCAGGC 395
 Db 402 TGAACCGCTGCTAGCTTCCATCAGGTGATTATACCTTTGACCGGAATGTGTATCCAGGC 461
 Qy 396 TGTAAATGAATGACAGAGAGATGCTGACCAATTCATTGAGCGCCACTCAGTCCCAAGT 455
 Db 462 TGTAAATGAATGACAGAGAGATGCTGACCAATTCATTGAGCGCCACTCAGTCCCAAGT 521
 Qy 456 CACATGAGCGGTTAATATGCTTGCATCATTTTTCAGATGTGAAATTTTGGCTGCT 515
 Db 522 CACATGAGCGGTTAATATGCTTGCATCATTTTTCAGATGTGAAATTTTGGCTGCT 581
 Qy 516 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATATGATGTATCAACA 575
 Db 582 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATATGATGTATCAACA 641
 Qy 576 AAT- -GTTGATGAGAGACATATGA 601
 Db 642 AATATGTTGATGAGAGACATATTA 669

RESULT 20
 AX869333
 LOCUS AX869333 816 bp DNA linear PAT 17-DEC-2003
 DEFINITION Sequence 4238 from Patent EPI074617.
 ACCESSION AX869333
 VERSION AX869333.1 GI:40024196
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Ota, T., Isozaki, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Makamatsu, A., Nagai, K. and Otsuki, T.
 TITLE Primers for synthesizing full-length cDNA and their use
 JOURNAL Patent: EP 1074617-A 4238 07-FEB-2001;
 Research Association for Biotechnology (JBP)
 FEATURES
 source 1. 816
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 91.1%; Score 547.6; DB 6; Length 816;
 Best Local Similarity 98.9%; Pred. No. 3.4e-126;
 Matches 562; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 95
 Db 102 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 161
 Qy 96 GTPAATATGCTGCCAATTCATGCGCACCTTAATATAGCAGACAGATGATGAGTGC 155
 Db 162 GTTAAATGCTGCCAATTCATGCGCACCTTAATATAGCAGACAGATGATGAGTGC 221
 Qy 156 TGGATAGCTCTACAGAGTACAGAGAGATGACCCAAAACAGAGAGAGAGAGAG 215
 Db 222 TGGATAGCTCTACAGAGTACAGAGAGATGACCCAAAACAGAGAGAGAGAGAG 281
 Qy 216 TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATCTTTATAGGAATATCAGT 275
 Db 282 TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATCTTTATAGGAATATCAGT 341
 Qy 276 TTTAATCAAGATGCTATGATGATGAGAAATTTAAGAGAAATTCATCAGCTGCTA 335
 Db 342 TTTAATCAAGATGCTATGATGATGAGAAATTTAAGAGAAATTCATCAGCTGCTA 401
 Qy 336 TGACCGTGTGCTTCCATCAGGTGATATACCTTTGACCGGAATGTGTATCCAGGC 395

Db 402 TGAACGCTGCTAGCTTCCATCAGGTGATTATACCTTTGACCGGAATGTGTATCCAGGC 461
 Qy 396 TGTAAATGAATGACAGAGAGATGCTGACCAATTCATTGAGCGCCACTCAGTCCCAAGT 455
 Db 462 TGTAAATGAATGACAGAGAGATGCTGACCAATTCATTGAGCGCCACTCAGTCCCAAGT 521
 Qy 456 CACATGAGCGGTTAATATGCTTGCATCATTTTTCAGATGTGAAATTTTGGCTGCT 515
 Db 522 CACATGAGCGGTTAATATGCTTGCATCATTTTTCAGATGTGAAATTTTGGCTGCT 581
 Qy 516 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATATGATGTATCAACA 575
 Db 582 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATATGATGTATCAACA 641
 Qy 576 AAT- -GTTGATGAGAGACATATGA 601
 Db 642 AATATGTTGATGAGAGACATATTA 669

RESULT 21
 AK097884
 LOCUS AK097884 1766 bp mRNA linear PRI 30-JAN-2004
 DEFINITION Homo sapiens cDNA FLJ40565 fis, clone THYMJ2004688, highly similar to Homo sapiens MDC-3.13 isoform 2 mRNA.
 ACCESSION AK097884
 VERSION AK097884.1 GI:21757783
 KEYWORDS Oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Saito, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oba, Y., Ishii, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosotani, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirakawa, S., Chiba, Y., Iehida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoshida, M., Hoshino, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T., Shihata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kungu, A., Iwakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobayake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Saeki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Maeno, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isozaki, T. and Sugano, S.
 TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs
 JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
 PUBMED 14702039

REFERENCE
 AUTHORS Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Mutsashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsumoto, K., Nakamura, Y., Sekine, M., Kikkawa, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A.,

Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 1766)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazuo-Kamata, Katsarazu, Chiba 292-0612, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction; Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.

FEATURES
 source
 1. 1766
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="THYM2004688"
 /tissue_type="thymus"
 /clone_id="THYM2"
 /note="cloning vector: pME18SFL3"

ORIGIN
 Query Match 83.3%; Score 500.8; DB 9; Length 1766;
 Best local Similarity 93.2%; Pred. No. 1.6e-114;
 Matches 561; Conservative 0; Mismatches 2; Indels 39; Gaps 2;

```

1  CTTCAAGGCTCCGGCGCGCGTGC-CGACTCTCCGATGCGCACAGATGCTTTAATTCGA 59
Db  141 CTTCAAGGCTCCGGCGCGCGTGC-CGACTCTCCGATGCGCACAGATGCTTTAATTCGA 200
QY  60 AAAAAGCTGGCGCTTTCAGGCAAAAAGAAATCTTGGTAAATGGTGTCCAAATCCATCG 119
Db  201 AAAAAGCTGGCGCTTTCAGGCAAAAAGAAATCTTGGTAAATGGTGTCCAAATCCATCG 260
QY  120 CCACACCTTAATAGACACACAGTAGTAGGTGTGTAGTGAAGCTTACAGATGCCA 179
Db  261 CCACACCTTAATAGACACACAGTAGTAGGTGTGTAGTGAAGCTTACAGATGCCA 309
QY  180 GGGAGTAGACCCAAAACAGAGAGAGAGAGAGATCATCAAGAACCTCATCAAGCAG 239
Db  310 -----CAGAGAAAGATCATCAAGAACCTCATCAAGCAG 342
QY  240 TCATCAAGCTGGCCATTCTTTAATAGAAATATCATGTTAATCAAGATGACCTAGCATGGA 299
Db  343 TCATCAAGCTGGCCATTCTTTAATAGAAATATCATGTTAATCAAGATGACCTAGCATGGA 402
QY  300 TGGAGAAATTTAGAGAAAGTTCATCAGCTTCTATGACCGGTGTCAAGTTTCCATCAGG 359
Db  403 TGGAGAAATTTAGAGAAAGTTCATCAGCTTCTATGACCGGTGTCAAGTTTCCATCAGG 462
QY  360 TGGATTATACCTTTGACCCGGAATGTGTATTCAGAGCTGTAAATGAATGACAGAGATGC 419
Db  463 TGGATTATACCTTTGACCCGGAATGTGTATTCAGAGCTGTAAATGAATGACAGAGATGC 522
QY  420 TGCACCAATCATTCAGCGCCACCTCATCGCCAAAGTCAATGACAGCGGTTAATATGCT 479
Db  523 TGCACCAATCATTCAGCGCCACCTCATCGCCAAAGTCAATGACAGCGGTTAATATGCT 582
QY  480 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATATCTTTGGGAATTTTA 539
Db  583 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATATCTTTGGGAATTTTA 642
QY  540 AACCCCACTTACAAAACCTATGTATGATGATCAACAAATGTGTGATGAGAGACATAT 599
Db  643 AACCCCACTTACAAAACCTATGTATGATGATCAACAAATGTGTGATGAGAGACATAT 702

```

QY 600 GA 601
 Db 703 GA 704

RESULT 22
 AC095257
 LOCUS
 DEFINITION
 AC095257
 AC095257.5 GI:24818022
 HTG: HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus (Norway rat)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 1 (bases 1 to 237561)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benamed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Gilly, M., Guerra, M., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, D.,
 Lorenz, L., Louised, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokelimeh, O., Okunodu, G., Olarinmoyegoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
 Plopper, F., Polindexter, A., Popovic, D., Prims, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tirney, A., Trejos, Z., Umanil, K.,
 Vales, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weiss, G., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 237561)
 Worley, K.C.
 TITLE
 JOURNAL
 Submitted (16-SEP-2001) Human Genome Sequencing Center, Department

REFERENCE
AUTHORS
TITLE
JOURNAL

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 237561)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 9, 2002 this sequence version replaced gi:23264563.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GDPL

Center clone name: CH230-10A12

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 229749 bases at least Q40

Consensus quality: 231209 bases at least Q30

Consensus quality: 232003 bases at least Q20

Estimated insert size: 232575; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 237561: contig of 237561 bp in length.

FEATURES

source

1. 237561

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-10A12"

1. 1156

/note="wgs end-extension

clone_end:Sp6"

6006..6586

/note="clone boundary

clone_end:Sp6

site:BCORI

end sequence:BH303075"

76320..79576

/note="wgs contig"

complement(232422..233289)

/note="clone boundary

clone_end:T7

site:BCORI

end sequence:BH303071"

234103..235440

/note="wgs_end-extension

clone_end:T7"

misc_feature 237561

/note="wgs_end-extension

clone_end:T7"

ORIGIN

Query Match 75.6%; Score 454.2; DB 2; Length 237561;

Best Local Similarity 85.9%; Pred. No. 9.4e-103;

Matches 504; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 15 CCCCGTCGCCACTCCTCCGATGGCCACAGATGCTTTAATTCGAAAACTGGCCGTTTC 74

DB 95025 CCTGTGTTTATTTCTCTAGTGGGTACAGATGCTTTAATTCGAAAACTGGCCGTTTC 95084

QY 75 AGCGCAAAAGAAAGATCTTGGTAAATGGTCCAAATCCATGGCCACCACTTAATAG 134

DB 95085 AGCGCAAAAGAAAGATCTTGGTAAATGGTCCAAATCCATGGCCACCACTTATG 95144

QY 135 ACAGACAAGTAGTAGGTGCTGATGAGCTTCAAGAGTGACAGAGGATACACCCAAA 194

DB 95145 ATGACACCGACGCGAGGCTCTGATGAGCTTCAAGGATGACCAAGGATACCCAGA 95204

QY 195 ACAAGAGAGGACGACAGAAATATCATAGAACTCATAGACATCATAGCTGGCCA 254

DB 95205 ACAAGAGAGGACGACAGAAATATCATAGAACTCATAGACATCATAGCTGGCCG 95264

QY 255 TTCTTATGAAATATCATAGTTAATCAAGATGAGCTAGCATTTGATGAGAAATTAAGA 314

DB 95265 TTCTTCAAGAGAAATATCATAGTTAATCAAGATGAGCTAGCATTTGATGAGAAATTAAGA 95324

QY 315 AGAAAGTTTCATGAGCTTGTATGACCTGTGTCAGTTTCATCAGGTGATTAATACCTTTC 374

DB 95325 GAAAGGTGACACGAGCTGGCATGACGATGATGAGCTTTCACAGGTAGATGATACCTTTC 95384

QY 375 ACCGGAATGTGTTATCCAGGCTGTTAATGATGACAGAGATGCTGACCAATCATTC 434

DB 95385 ACCGGAATGTGTTATCCAGGCTGTTAATGATGACAGAGATGCTGACCAATCATTC 95444

QY 435 AGCGCAACCTCATGTCGCAAGTACATGACGCGGTTAATATGCTTTGATCATTTTTCAG 494

DB 95445 AGCGCAACCTCATGTCGCAAGTACATGACGCGGTTAATATGCTTTGATCATTTTTCAG 95504

QY 495 ATTTGAAATTTTGGCTGCTGCTGTTAATATCCTTTTGGAAATTTAAACCCACTTACAA 554

DB 95505 ATTTGAAATTTTGGCTGCTGCTGTTAATATCCTTTTGGAAATTTAAACCCACTTACAA 95564

QY 555 AACTATGTGATGCTATCAACAAATGTTGATGAAGAACAATATGA 601

DB 95565 AACTATGTGATGCTATCAACAAATGTTGATGAAGAACAATATGA 95611

RESULT 23

BC009090

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

MUS

REFERENCE

AUTHORS

1 (bases 1 to 1728)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Strusberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buehler,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshitaki,S.,

Carninci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullan,S.J., Bosak,S.A., McMan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

1728 bp mRNA linear ROD 30-UTN-2004

Mus musculus tumor necrosis factor, alpha-induced protein 8, mRNA

(CDNA clone MGC:11714 IMAGE:396593), complete cds.

Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Rhee, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL
 PUBLISHED
 REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 JOURNAL

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CNA Library Preparation: Life Technologies, Inc.
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsgaard, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nambayati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 17 Row: n Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19527361.

FEATURES
 source
 1. 1728
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:11714 IMAGE:3965693"
 /tissue="Mammary tumor. Metallothionin-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
 /clone_1lb="NCI CGAP_Mam1"
 /lab_host="DHI0B"
 /note="Vector: PCMV-SPORT6"
 1. 1728
 /gene="Tnfrif8"
 /note="synonyms: Nded, Gg2-1, Sec-2"
 /db_xref="locusid:106869"
 /db_xref="MGI:2147191"
 70. 666
 /gene="Tnfrif8"
 /codon_start=1
 /product="TNF-induced protein"
 /protein_id="AAH09090.1"
 /db_xref="GI:14318586"
 /db_xref="locusid:106869"
 /db_xref="MGI:2147191"
 /translation="MLSEAEPREVATDVNSKNLAVQOKILGKRVSKATITLID
 DTSEVLDLRYVTKYONKKEARVYIKLITVYIKLAVLHNNQFNDELAIMKRF
 KKKVHOLAMTVUSFHOVEYFPDNYSTRLLNCRRLHETIQHNLAKSHGRVNVFD
 HSDCPFLAALNIPGKFKPHQLCDGINKMLDEENT"

ORIGIN
 Query Match 73.9%; Score 444.4; DB 10; Length 1728;
 Best Local Similarity 86.6%; Pred. No. 2,3e-100;
 Matches 490; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGGCAAGAAAGATCTTGG 95
 Db 101 TGGCTCAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGGCAAGAAAGATCTTGG 160
 QY 96 GTAAATGGGTGCAATTCATGCGACCACTTAATGAGCGACAAAGTAGAGGTGC 155
 Db 161 GCAAAATGGTATCAATTCATGCGACCACTTAATGAGCGACAAAGTAGAGGTGC 220
 QY 156 TGGATGAGCTCTACAGATGACGAGGAGTACACCCAAAACAAGAGAGCGAGAGA 215
 Db 221 TAGATGAGCTGTACAGGAGTACCAAGAGTACACCCAAAACAAGAGAGCGAGAGG 280
 QY 216 TCATCAAGAACTCATCAAGACGATCATCAAGCTGCATCTTTATAGGAATATCAGT 275
 Db 281 TCATCAAGAACTCATCAAGACGATCATCAAGCTGCATCTTCACAGGAACATCAGT 340
 QY 276 TTTATCAAGATGAGCTGATTCATGATGAGAAATTTAAGAAAGATTCACACTTGTCTA 335
 Db 341 TCATCAAGACGAGCTGCGCTCATGAGAAATTTAAGAAAGATTCACACTTGTCCA 400
 QY 336 TGACCGTGTGCTGATTCATCAGTGGATTAATCTTTGACCGGAATGTGTATCCAGG 395
 Db 401 TGACGCTGTGCTGATTCATCAGTGGATTAATCTTTGACCGGAATGTGTATCCAGG 460
 QY 396 TGTAAATGATGATGACAGAGATGCTGACCAATTCATGAGCGCACTTACGCAAGT 455
 Db 461 TGCTGAACGATGCTGACAGATGCTTCAAGAGATTCATGAGCGCACTTACGCAAGT 520
 QY 456 CACATGAGCGGTTAATGATGCTTGTGATTCATTTTCAATTTGATTTTGGCTGCT 515
 Db 521 CTCACGAGCGGTTAATGATGCTTGTGATTCATTTTCAATTTTGGCTGCT 580
 QY 516 TGTATATCTCTTTGGAAATTTAAACCCACTTACAAAACATGATGATCAACA 575
 Db 581 TGTATATCTCTTTGGAAATTTAAACCCACTTACAAAACATGATGATCAACA 640
 QY 576 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
 Db 641 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666

RESULT 24
 AC120859 184327 bp DNA linear ROD 05-AUG-2004
 LOCUS Mus musculus chromosome 18, clone RP24-200D2, complete sequence.
 DEFINITION AC120859
 ACCESSION AC120859
 VERSION AC120859.12 GI:50979395
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 (bases 1 to 184327)
 Birren, B., Nussbaum, C. and Lander, E.
 Unpublished
 2 (bases 1 to 184327)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouknight, B., Brown, A., Camarero, J., Campiano, A., Chang, J., Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeBartolo, R., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardy, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kages, A., Karacas, A., Kells, C., Lakoque, K., Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, J., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

TITLE
JOURNAL
REFERENCE
AUTHORS

Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trifilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 184327)
 Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearlano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Punthang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (22-JUN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 184327)
 Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearlano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Punthang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (05-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 5, 2004 this sequence version replaced gi:49035085.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@road.mit.edu
 Project Information
 Center project name: L21053
 Center clone name: 200_D_2

FEATURES

source	Location/Qualifiers
misc_feature	1..184327 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="18" /map="18" /clone="RP24-200D2" /clone_lib="RPCT-24 Male Mouse BAC"
repeat_region	1..4 /note="clone_boundary clone_end:SP6 site:MboI"
repeat_region	675..700 /rpt_family="AT_rich"
repeat_region	984..1007 /rpt_family="AT_rich"
repeat_region	3199..3484 /rpt_family="B4A"
repeat_region	3506..3534 /rpt_family="CA)n" complement(4228..4578)
repeat_region	4848..5049 /rpt_family="B3"
repeat_region	5526..5572 /rpt_family="AT_rich"
repeat_region	6521..6815 /rpt_family="Lx9"
repeat_region	6890..6921 /rpt_family="AT_rich"
repeat_region	8693..8735 /rpt_family="TTCC)n"
repeat_region	8699..8700 /note="probably CC"
repeat_region	9208..9234 /rpt_family="TTTG)n" complement(9236..9335)
repeat_region	complement(9603..9751) /rpt_family="B1F"
repeat_region	10571..10599 /rpt_family="AT_rich"
repeat_region	10637..10660 /rpt_family="AT_rich"
repeat_region	10951..11146 /rpt_family="B2_Mm1"
repeat_region	12671..12730 /rpt_family="TG)n"
repeat_region	13141..13214 /rpt_family="ID3"
repeat_region	13228..13256 /rpt_family="PB1D10"
repeat_region	13257..13323 /rpt_family="ID_B1"
repeat_region	13331..13358 /rpt_family="AT_rich"
repeat_region	13483..13667 /rpt_family="B3"
repeat_region	13769..14378 /rpt_family="LIM4"
repeat_region	14519..14619 /rpt_family="L1_MM"
repeat_region	complement(14772..15069) /rpt_family="L1_MM"
repeat_region	15389..15420 /rpt_family="AT_rich"
repeat_region	17608..17670 /rpt_family="TATATG)n"
repeat_region	17773..17801 /rpt_family="AT_rich"
repeat_region	17825..17850 /rpt_family="AT_rich"


```

repeat_region      18110..18163
                    /rpt_family="AT_rich"
repeat_region      20210..20731
                    /rpt_family="L1_MM"
repeat_region      20759..20876
                    /rpt_family="RSINE1"
repeat_region      21592..21799
                    /rpt_family="B1_MM"
repeat_region      21818..21837
                    /rpt_family="(A)n"
repeat_region      21916..21949
                    /rpt_family="AT_rich"
repeat_region      22545..22831
                    /rpt_family="Lx7"
repeat_region      22892..23011
                    /rpt_family="B1_MM"
repeat_region      22949..22974
                    /note="single clone coverage"
repeat_region      23095..23294
                    /rpt_family="Lx7"
repeat_region      23912..24049
                    /rpt_family="B1_MM"
repeat_region      24017..24236
                    /rpt_family="MTC"
repeat_region      24238..24355
                    /rpt_family="MTC"

Query Match      73.9%; Score 444.4; DB 10; Length 184327;
Best Local Similarity 85.9%; Pred. No. 2.6e-100;
Matches 493; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

28  TCTCCCATGCGCCACAGATGCTTTATTCGCAAAACCTGCGCTTATGAGCACAAGAA 87
144846 TTCTGCGATGAGCTACAGATGCTTCAATTCGCAAAACCTGCGCTTATGAGCACAAGAA 144905

88  GATCTTGAGTAAATGCTGCTCAAAATCCATGCGCCACCTTAATAGACACAAGTAG 147
144906 GATCTTGAGTAAATGCTGCTCAAAATCCATGCGCCACCTTAATAGACACAAGTAG 144965

148  TGAAGTCTGAGTACGCTTCAAGAGTGACCAAGGATACACCCAAAACAAGAGAGC 207
144966 CGAGTCTGAGTACGCTTCAAGAGTGACCAAGGATACACCCAAAACAAGAGAGC 145025

208  AGAAGATGATCAAGAACTCTCAAGACAGTCAATCAAGCTGCTTTTATAGAA 267
145026 GGAAGAGTCAAGAACTCTCAAGACAGTCAATCAAGCTGCTTTTATAGAA 145085

268  TAATCAAGTCAATCAAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 327
145086 CAATCAAGTCAATCAAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 145145

328  GCTTGCTATGACCGTGTGCTCAATGCTTCAATGCTGATTAATCTTTCGAGGATGTGT 387
145146 GCTTGCTATGACCGTGTGCTCAATGCTTCAATGCTGATTAATCTTTCGAGGATGTGT 145205

388  ATCCAGGCTTTTAATGATGACAGAGATGCTGACCAAAATCAATTCAGGCGCACTGAC 447
145206 GTCAGGCTGCTGATGACAGAGATGCTGACCAAAATCAATTCAGGCGCACTGAC 145265

448  TGCAAGTCAATGACAGGCTGATTAATGCTTTCATGCTTTCATGCTTTCATGCTTTCAT 507
145266 CGCAAGTCAATGACAGGCTGATTAATGCTTTCATGCTTTCATGCTTTCATGCTTTCAT 145325

508  GAGTCTGCTTATATCTTTTGGGATTTTAAACCCCACTTACAAAACCTATGTGATG 567
145326 GAGTCTGCTTATATCTTTTGGGATTTTAAACCCCACTTACAAAACCTATGTGATG 145385

568  TATCACAAAATGTTGATGAGAGCAATATGA 601
145386 TATCACAAAATGTTGATGAGAGCAATATGA 145419

```

```

LOCUS      AJ720906      2009 bp      mRNA      linear      VRT 30-SEP-2004
DEFINITION Gallus gallus mRNA for hypothetical protein, clone 29h8.
ACCESSION  AJ720906
VERSION    AJ720906.1 GI:53136471
KEYWORDS   ORF1.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1
AUTHORS   Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezubov, Y., Zaim, J.,
            Friedler, P., Kuter, S., Biagodatski, A., Kostowska, D., Koter, M.,
            Plachy, O., Carninci, P., Hayashizaki, Y. and Buerstedde, J.M.
TITLE      Full-length cDNAs from bursa lymphocytes to facilitate gene
            function analysis
JOURNAL    Unpublished
            2 (bases 1 to 2009)
AUTHORS    Caldwell, R.B.
REFERENCE  Direct Submission
            Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
            Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
            1, D-85764 Neuherberg, GERMANY
FEATURES   location/Qualifiers
            source
            1..2009
            /organism="Gallus gallus"
            /mol_type="mRNA"
            /strain="CB"
            /db_xref="taxon:9031"
            /clone="29h8"
            /cell_type="lymphocyte"
            /tissue_type="bursa"
            /clone_lib="rikeni"
            /dev_stage="2 weeks old"
            1..2009
            /locus_tag="RCJMB04_29h8"
            <1..188
            /locus_tag="RCJMB04_29h8"
            189..755
            /locus_tag="RCJMB04_29h8"
            /note="ORF1"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="CA32565.1"
            /db_xref="GI:53136472"
            /translation="MATDVFNKSLAIQAKKILGKMSKSIATTLIDPSSPVLDEL
            YRVTKEVYONKKEAEKIKNLIKIVLKALIVRNQFNQDEILAMKPKKHVQLAKT
            VVSHQVDTTPRPPLSKLNDRELIHQIRHLTKSHGRVNVVFDHPSDEPLAA
            LYNPSPYKHLQKLCDSGVNRMIDBSNT"
            756..>2009
            /locus_tag="RCJMB04_29h8"

3' UTR

ORIGIN
Query Match      68.0%; Score 408.6; DB 5; Length 2009;
Best Local Similarity 82.5%; Pred. No. 2.1e-91;
Matches 468; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

33  CGATGGCAAGATGCTTTAATTCAAAACCTGCGCTTACGACAAAAGAAATGTC 92
187 CAATGGCAAGATGCTTTCAATTCAAAAGCTTTCAGGCCCAAGAAAGATTC 246

93  TGGGTAATATGCTGTCCAATCCATGCGACCACTTAATAGACAGACAAGTAGAG 152
247 TTGGGAAATATGATCAAGTCAATGCAACTCTTGAATAGATATACAGAGTAGT 306

153  TGGTGAATGAGCTTCAAGAGTACCAAGGAGTACCCAAAACAAGAGAGGAGAGA 212
307 TTTAATATAGCTTCAAGAGTACCAAGGAGTACCGAAAACAAGAGAGAGAGA 366

213  AGATCATCAAGAACTCATCAAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 272
367 AGATCATTAATAACCTCATTAATAATGCTCTCAAAATTCGCAATTCCTACAGAACATC 426

```

Qy	273	GGTTAAACAGATGAGCTACATGTATGAGAAATTTAAGAAGAAAGTTATCATCAGCTTG	332
Db	427	AATTTAAACAAAGATGAATAATTCACATCGATGGAGAAATTCAAGAAAGAAATTCATCAGCTGG	486
Qy	333	CTATGACCGTGTCAGTTCCATGCATGAGTGAGATTATACCTTTGACCGGAATGTTATCA	392
Db	487	CGAAGACGGTGTCTCAGTTTCCATCAGGTGAGATTATACCTTTGACGGAATTTCTTGTC	546
Qy	393	GGCTGTTAAATGAATGACAGAGATGTCGACCAATCATTCAGGCCACCTCACTGCGCA	452
Db	547	AACTGTTGAATGACGTGAGAGAGCTGCTTCATCAGATCATTCAGGTCACCTAACTGCGA	606
Qy	453	AGTCACATGACGGGTATTAATATGTCCTTGATCATTTTTCAGATTGTGAATTTTGGCTG	512
Db	607	AATGCGATGACGATGTCAACATGHTGTTGATCACTTCTCGATTGTGAATTTTGGCTG	666
Qy	513	CTTGTTAATATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGTGATGTC	572
Db	667	CCTTGTTAATATCCTTTTGGACCTTATTAACCTTCATTTCAAAAAATTGTGTATGTGTCA	726
Qy	573	ACAAATGTGATGATGAGAAACATAT	599
Db	727	ACAGATGTCTAGATGAGGGAACATAT	753

RESULT	26
LOCUS	BDI49633
DEFINITION	BDI49633 587 bp DNA linear PAT 17-JAN-2003
ACCESSION	BDI49633
VERSION	BDI49633.1 GI:27855391
KEYWORDS	JP 2002191363-A/4476.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 587)
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T. Primer for synthesizing full-length cDNA and use thereof Patent: JP 2002191363-A 4476 09-JUL-2002;
TITLE	HELIX RESEARCH INSTITUTE
JOURNAL	OS Homo sapiens (human) PN JP 2002191363-A/4476 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI JUNICHI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI MAGAI,TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10,
COMMENT	PC C12P21/02,C12Q1/66//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT source 1..587 FT location/Qualifiers 1..587 location/Qualifiers 1..587 /organism='Homo sapiens (human)'. /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
FEATURES	
SOURCE	
ORIGIN	

	Query Match	Similarity	Score	DB #	Length
Best Local	96.6%	Pred. No. 6.5e-8;			
Matches	422;	Conservative	0;	Mismatches 13;	Indels 2; Gaps 2;
Qy	1	CTTCAGCGCTCCGCGCGCGCGCTCGC	-GACATCTCTCCGATGAGCCACAGATGCTTTAATTATTC	CA	59
Db	105	CTTCAGCGCTCCGCGCGCGCGCTCGC	CACTCTCCGATGAGCCACAGATGCTTTAATTATTC	CA	164

QY	60	AAAACTGGCGGTCAGGACAAAAAGAAAGATCTTGGGTMAAATGGTGTCCAAATCCATCG	119
Db	165	AAAACCTTGCGCGTTCAGGACAAAAAGAAAGATCTTGGGTMAAATGGTGTCCAAATCCATCG	224
QY	120	CCACCACCTTAATAGACAGACAAAGTAGTAGAGTCTGGATGAGCTCTACAGATGACCA	179
Db	225	CCACCACCTTAATAGACAGACAAAGTAGTAGAGTCTGGATGAGCTCTACAGATGACCA	284
QY	180	GGGAGTACACCCAAAAACAAGAGGCGAGAGAAAGATCATCAAGAACTCATCAAGACAG	239
Db	285	GGGAGTACACCCAAAAACAAGAGGCGAGAGAAAGATCATCAAGAACTCATCAAGACAG	344
QY	240	TCATCAAGCTGGCCATCTTTTATAGGAATAATCAGTTAATCAAGATGAGCTAGCATTTGA	299
Db	345	TCATCAAGCTGGCCATCTTTTATAGGAATAATCAGTTAATCAAGATGAGCTAGCATTTGA	404
QY	300	TGGAGAAATTTAAGAGAAGATTCATCAGCTTGCTATACCGGTGTCAGTTCCATCAGG	359
Db	405	TGGAGAAATTTAAGAGAAGATTCATCAGCTTGCTATACCGGTGTCAGTTCCATCAGG	464
QY	360	TGATTTATACCTTTGAGCCGGAATGTGTTATCCAGGCTGTAAATGAATGACAG-AGAGATG	418
Db	465	GGGGTTATACCTTTGAGCCGGAATGTGTTATCCAGGCTGTAAATGAATGACANAAGAAATG	524
QY	419	CTGCACCAATCATCTTCA	435
Db	525	CTGCACCAAAACANTTA	541

RESULT 27				
AX869571				
LOCUS				
DEFINITION	AX869571	587 bp	DNA	linear
ACCESSION	AX869571			
VERSION	AX869571.1			
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1			
JOURNAL	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T. Primeres for synthesizing full-length cDNA and their use Patent: EP 1074617-A 4476 07-FEB-2001; Research Association For Biotechnology (JP)			
FEATURES	location/Qualifiers			
source	1..587			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			

Query Match	Similarity	65.7%	Score 394.6	DB 6	Length 587
Beet Local	Similarity	96.6%	Pred. No. 6.5e-8		
Matches	Conservative	0	Mismatches	13	Indels
				2	Gaps
QY	1	CTTCAGCCCTCCGGCGCCGTCGC--CGACTCTCCGCATGGCCACAGATGCTTTAAATTC	59		
DB	105	CTTAGCGCTCCGGCGCGCGCCGCACTCCGATCGGCGCACAGATGCTTTAAATTC	164		
QY	60	AAAACCTGGCCGTTAGGCACAAAAGAAAGATCTTGGGTAAATGGTGTCCAAATTCATG	119		
DB	165	AAAACCTGGCCGTTAGGCACAAAAGAAAGATCTTGGGTAAATGGTGTCCAAATTCATG	224		
QY	120	CCACCACCTTAAATAGACGACAAATAGTAGAGTCTGATGAGCTCTACAGATGACCA	179		
DB	225	CCACCACCTTAAATAGACGACAAATAGTAGAGTCTGATGAGCTCTACAGATGACCA	284		
QY	180	GGAGTAGACCCAAAACAAAGAGGAGGACAGAGATCATCAAGAACTTCATCAAGACG	239		
DB	285	GGGAGTAGACCCAAAACAAAGAGGAGGACAGAGATCATCAAGAACTTCATCAAGACG	344		

Query Match	59.2%	Score 355.6	DB 5	Length 1559
Best Local Similarity	76.1%	Pred. No. 3.9e-78		
Matches 452	Conservative 0	Mismatches 139	Indels 3	Gaps 1
ORIGIN				
FEATURES				
Source				
1..1559				
/organism="Xenopus tropicalis"				
/mol_type="mRNA"				
/db_xref="taxon:8364"				
/clone="UTBg053008"				
/clone_lib="XGC-egg"				
/dev_stage="egg"				
Query Match	59.2%	Score 355.6	DB 5	Length 1559
Best Local Similarity	76.1%	Pred. No. 3.9e-78		
Matches 452	Conservative 0	Mismatches 139	Indels 3	Gaps 1
ORIGIN				
FEATURES				
Source				
1..1559				
/organism="Xenopus tropicalis"				
/mol_type="mRNA"				
/db_xref="taxon:8364"				
/clone="UTBg053008"				
/clone_lib="XGC-egg"				
/dev_stage="egg"				

QY	248	CTGGCAGTCTCTTTATAGGAATAATACAGTTTAAATGAATAGAGCTAGATTTGATGGAGAA	307
Db	371	CTGGCAGTCTCTTACAGGAATAATACAGTTTAAATGAAGAGATTTGACCTATGGAGAG	430
QY	308	TTTATAGAGAAAGTTTCATCAGCTTGCTCTATGACCGCTGGTCAGTTTCCATCAGGTGATTTAT	367
Db	431	TTCAAGAGAAAGTTCATCAGCTTGCCATGAGCTGTGGTCAAGCTTTTATCAGGTGGAAATAT	490
QY	368	ACCTTTGACCGGAATGTGTTATCCAGGCTGTTAATGAATGACAGAGAGATGCTGCACCA	427
Db	491	ACCTTTGACCGGAACGAGCTTTCCAAATTTGTTAATGATGATGCGCGAGAGCTGCTCCATCAG	550
QY	428	ATCATTTAGGCGCCACCTCAGCTGCCAATGACATGACAGGAGTATTAATGCTTTGATCAT	487
Db	551	GTCATACAGCGCCATCTCAGCGCAAAAGTCTCAGAGCGTGTAAATACGATTTGACAC	610
QY	488	TTTTTCAGATTTGTGAATTTTGGCTGCTGCTGTATATCTTTTGGGAATTTTAAACCCAC	547
Db	611	TTTTCAATATGTGAATTTTGGCTGCTGCTGTACATCCCTTTGACCTTTCAACAGAC	670
QY	548	TTACAAAACCTATGTGATGCTATGACACAAAATGTTGATGAAGAAACATATGA	601
Db	671	CTTCAGAGAGCTCTGTGAAGAGTCAACAAAATGCTGACGAGACGAATATTTTAA	724
RESULT 29			
BC072904			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			

REMARK
COMMENT

Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Igor David
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Telia Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliq Scherch, Ursula Skalska, Duane Smalls, Jeff Scott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 152 Row: J Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1. 1193

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="MGC:80354 IMAGE:5074412"

/tissue_type="Ovary, Xenopus"

/clone_lib="NICH_XGC_Ov1"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 1193

/gene="MGC80354"

/db_xref="LocusID:443963"

153. 722

/gene="MGC80354"

/codon_start=1

/product="MGC80354 protein"

/protein_id="AAH72904.1"

/db_xref="GI:49115086"

/LocusID:443963"

/translation="MATDIFNSKTLAVQAKILGMASSTYATSLIDTSGEVLDE
LYQVREYQNKDAEKITKILYIKLAVLRNNQFNEEIEGLMEKRFKRVHQIAM
TVVSFQVETTFDRNVLSKLNECRELHLOVIRHLTAQSHGRVNVVDFHSCFELA
ALVNPFGPKKHLQRLCNGVNMKLDENI"

ORIGIN

Query Match 58.9%; Score 354; DB 5; Length 1193;
Best Local Similarity 75.9%; Pred. No. 9.7e-78;

Matches 451; Conservative 0; Mismatches 140; Indels 3; Gaps 1;

11 CCGGCGCCGTCGCGACTCTCCGATGCGCAGATGCTTTAATTCAAAACCTGGCC 70

129 CCGGCGCCGTCGCGACTCTCCGATGCGCAGATGCTTTAATTCAAAACCTGGCC 188

71 GTTCAGGACAAAGAAAGATCTTGGGTAAATG---TGTCCAAATCCATCGCACACC 127

189 GTTCAAGCCCAAGAAAGATCTTGGCAAGATGCTTCAATTAATCATGACAGCTCC 248

128 TTAATAGACACAAAGTGTGCTGATGAGCTTACAGAGTGAACAGGAGTAC 187

249 CTGATGATGATCAAGTGGCGAAGTTCTAGATGAATCTACCAAGTACTAGGAGTAC 308

188 ACCCAAAACAAAGAGGCGAAGATATCAAGAACTCTATCAAGACATCATCAAG 247

309 ACCCAAAACAAAGAGTGGCGAAGAAATCAAAAGAACTCTCAAAACATCATCAAG 368

QY 248 CTGGCATCTCTTATAGAAATATCATGATTATCAAGATGAGTACATGATGAGAAA 307

DB 369 CTGGCAGCTTTATACCGAAACAGATTAAACGAGAGAGATTGCTTTATGAGAA 428

QY 308 TTTAAGAAAGATTCATCAGCTTCTATGACCGTGTGATGTTTCATCAGGTGATAT 367

DB 429 TTTAAGAAAGATTCATCAGCTTCTATGACCGTGTGATGTTTCATCAGGTGATAT 488

QY 368 ACCCTTGACCGGAAATGTGTATTCAGAGCTGTTAAATGATGAGAGATGTGACCAA 427

DB 489 ACCCTTGACCGGAAATGTGTATTCAGAGCTGTTAAATGATGAGAGATGTGACCAA 548

QY 428 ATCATTCAGCGCACCTCATCTCCAGTACATGACGCGGTAAATATGCTTTATCAT 487

DB 549 GTCATTCAGCGCACCTCATCTCCAGTACATGACGCGGTAAATATGCTTTATCAT 608

QY 488 TTTTCAGATTTGAAATTTTGGCTGCTTTGTAATTCCTTTGGAAATTTAAACCCAC 547

DB 609 TTTTCAGATTTGAAATTTTGGCTGCTTTGTAATTCCTTTGGAAATTTAAACCCAC 668

QY 548 TTTCAAAACCTATGATGATGATCAACAAATGTTGATGAGAGACATATGA 601

DB 669 CTTCAGAGACTCTGTACAGAGTCAACAAATGCTGATGAGAGACATATTTAA 722

RESULT 30
AF098933
LOCUS
DEFINITION
Homo sapiens head and neck tumor and metastasis related protein
mRNA, partial cds.

AF098933

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

AF098933.1 GI:6851132

```

Db      1 GATGAGAAATTTAAGAAAGATTCAATGCTGTATGACCGTGGTCACTTCATCA 60
QY      358 GGTGATTAATACCTTTGACCGGAAATGTTATTCAGCGCTGTTAAATGACAGAGAT 417
Db      61 GGTGATTAATACCTTTGACCGGAAATGTTATTCAGCGCTGTTAAATGACAGAGAT 120
QY      418 GGTGACCAATATCTTACAGCGCCCTCACGCGCAATGATCATGAGAGGGTATATGT 477
Db      121 GGTGACCAATATCTTACAGCGCCCTCACGCGCAATGATCATGAGAGGGTATATGT 180
QY      478 CTTGATCATTTTTCAGATTTGATTTGGCTGCTGTATATCTTTGGGAATTT 537
Db      181 CTTGATCATTTTTCAGATTTGATTTGGCTGCTGTATATCTTTGGGAATTT 240
QY      538 TAAACCCACTTACAAAACTATGTATGATGATCAACAAATGTTGATGAAGAAT 597
Db      241 TAAACCCACTTACAAAACTATGTATGATGATCAACAAATGTTGATGAAGAAT 300
QY      598 ATGA 601
Db      301 ATGA 304

RESULT 31
AC118938 164258 bp DNA linear HTG 24-FEB-2003
LOCUS Mus musculus clone RP24-174015, WORKING DRAFT SEQUENCE, 8 unordered
DEFINITION pieces.
AC118938 3 GI:28475940
ACCESSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 164258)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP24-174015
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 164258)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Chepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Fairo,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karacas,A., Kells,C., Lacombe,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,U., Peterson,K., Phunhahang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strausen,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
Submitted (21-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164258)
REFERENCE
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArillano,K.,
Diaz,J.S., Dodge,S., Doolley,K., Dorris,L., Erickson,J., Fairo,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hates,N., Hagopian,D., Hagos,B.,

```

```

Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karacas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,U., Peterson,K., Phunhahang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Stubs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
Submitted (24-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 24, 2003 this sequence version replaced gi:28261544.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25165
Center clone name: 174.O.15
----- Summary Statistics
Sequencing vector: Plasmid, n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162606 bases at least Q40
Consensus quality: 162965 bases at least Q30
Consensus quality: 163183 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 163558; sum-of-coverage
Quality coverage: 11.0 in Q20 bases; agarose-fp
Quality coverage: 11.0 in Q20 bases; sum-of-coverage
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 29707: contig of 29707 bp in length
29708 29807: gap of 100 bp
29808 32340: contig of 2533 bp in length
32341 32440: gap of 100 bp
32441 37689: contig of 5248 bp in length
37690 37789: gap of 100 bp
37789 104529: contig of 66740 bp in length
104530 104629: gap of 100 bp
104630 115026: contig of 10397 bp in length
115027 115126: gap of 100 bp
115127 131677: contig of 16551 bp in length
131678 131777: gap of 100 bp
131778 158793: contig of 27016 bp in length
158794 158893: gap of 100 bp
158894 164258: contig of 5365 bp in length.
Location/Qualifiers
1.164258
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-174015"
/clone_1ib="RP24-174015"
1.29707
/notes="assembly_fragment
clone_end:536

```

FEATURES

source

misc_feature

```

misc_feature      29808..32340
                  /note="assembly_fragment"
misc_feature      32441..37689
                  /note="assembly_fragment"
misc_feature      37790..104529
                  /note="assembly_fragment"
misc_feature      104630..115026
                  /note="assembly_fragment"
misc_feature      115127..133677
                  /note="assembly_fragment"
misc_feature      131778..158793
                  /note="assembly_fragment"
misc_feature      158894..164258
                  /note="assembly_fragment"
                  clone_end:T7
                  vector_side:right"

```

ORIGIN

Query Match 42.8%; Score 257.4; DB 2; Length 164258;

Best Local Similarity 65.7%; Pred. No. 1.7e-53;

Matches 375; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

```

QY 31 TCCGATGCCACAGATGCTTTAATTCAAAAAAGCTGGCGTTCAGGCAAAAAAGAT 90
DB 88007 TACCATAGGTCCTCATGTGTAGTTCAAGAACTCGCCCTCAAGCCAGAAAGAT 87948
QY 91 CTGGGTAATAATGTTGTCATTCACCAATCCGACACCTTAATAGACACCAAGTGTGA 150
DB 87947 CCTGACCAAGATGACCAAGCAAACTGTGGCAACATGATTTGATGACACCAAGTGA 87888
QY 151 GGTGTCGATGAGTCTTACAGATGACCAAGGAGTACCCCAAAAGAGAGGACAGA 210
DB 87887 GATCTTTGATGAGCTTCAAAATCACAAAATACACCCCAAAAGAGAGGACCCA 87828
QY 211 GAAGATCATCAAGAACCTCATCAAGACAGTCAATCAAGTGGCCATTCTTTATAGAAATA 270
DB 87827 CAAGATCATCAAGAGAGTCTATCAAGGTGGCAATCAAAATGGCATTCTTACCGAACA 87768
QY 271 TCAGTTTATCAAGAGTCTTACAGATGAGTGAAGAAATTAAGAAAGTTCATCATGCT 330
DB 87767 ACAGTTCAGTCAAGAGAGGATTAATTTGGGAAATCCGGAAGAACTGAACAGAC 87708
QY 331 TGCATGACCGTGTGATGCTTCCATCAGGTGATTAATCTTTCAGCGGAATGTGTATC 390
DB 87707 TCGATGACCATGATGATGCTTCAAGATGAGTATCTTTCATGATGATGATGCTTCT 87648
QY 391 CAGGCTGTTAATGAATGACAGAGATGCTGCAACCAATCAATCAGCGCCACTCATGTC 450
DB 87647 TAAGCTTTCGATGATGAGAGAGCTGATCAATGAATGATGATGATGATGATGATG 87588
QY 451 CAAGTCAATGACCGGTTAATTAATGCTTTCATGATGATGATGATGATGATGATGATG 510
DB 87587 CAGATCCATGAGCGATTAACATCTTCAACCACTTTCGATGATGATGATGATGATG 87528
QY 511 TGCCTGTTAATTCCTTTTGGGATTTTAAACCCCACTTAACAAACCTAATGATGATGAT 570
DB 87527 CACTCTCTAATGCTGCGCATGGAACCTGACGCGCCATCTCAAGAGATTTGCAAGAAAT 87468
QY 571 CAACAATAATGTTGATGAGAGAGAACTATGA 601
DB 87467 CAATAAATGTTAGACGACAGATCTCTGA 87437

```

RESULT 32

```

AC107848          195574 bp      DNA      linear      HTG 22-MAR-2004
LOCUS             Mus musculus chromosome 9 clone RP23-321A19 map 9, *** SEQUENCING
AC107848          AC107848
ACCESSION          AC107848.5 GI:45598920
VERSION            HTG; HTGS_PHASE1; HTGS_FUILLTOP; HTGS_ACTIVEPIN.
KEYWORDS            Mus musculus (house mouse)
SOURCE

```

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

1 (bases 1 to 195574)

Birren,B., Nusbaum,C. and Lander,E.

Mus musculus chromosome 9, clone RP23-321A19

Unpublished

2 (bases 1 to 195574)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,

Brown,A., Camarata,U., Campopiano,A., Chang,J., Chazaro,B.,

Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,

Cooke,P., Dearliano,K., Dewar,K., Diaz,J.S., Dodes,S., Faro,S.,

Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardys,S.,

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,

Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M.,

McEwan,P., McKernan,K., Meidrim,J., Menus,L., Mihova,T.,

Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,

Norman,K.H., O'Connor,T., O'Donnell,E., O'Neill,D., Oliver,J.,

Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,

Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,

Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,

Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strause,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Triggillo,J., Vasilev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.D., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 195574)

Birren,B., Nusbaum,C., Lander,E., Aboueleil,A., Allen,N.,

Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,

Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chapel,Y.,

Collymore,A., Cook,A., Cooke,P., Corum,B., Dearliano,K.,

Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,

Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardys,S.,

Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,

Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,

Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,

Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,

Melirim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,

Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,E.,

O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,

Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,

Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,

Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,

Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,

Vasilev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,

Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (22-MAR-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 22, 2004 this sequence version replaced gi:44886703.

All repeats were identified using RepeatMasker:

Smit,A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Center project name: 120615

Center clone name: 321_A_19

* NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. The true order of the pieces

* is not known and their order in this sequence record is

27

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRRAK Plate: 101 Row: C Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 41056158.
Location/Qualifiers
1. 3986
/organism="Danio rerio"
/mol_type="mRNA"
/strain="AB"
/db_xref="taxon:7955"
/clone="MGC:55331 IMAGE:2601021"
/tissue_type="whole body, adult, (one male and one female, including unfertilized eggs)"
/clone_11b="Sugano Kawakami zebrafish DRA"


```

/lab host="DH10B"
/notg="Vector: pME18S-FL3"
1..3986
/gene="tnfaip8"
/notes="synonyms: MGC55331, zgc:55331"
/db_xref="LocustID:393303"
196..756
/gene="tnfaip8"
/codon_start=1
/product="tumor necrosis factor, alpha-induced protein 8"
/protein_id="AAH52765.1"
/db_xref="GI:31127058"
/db_xref="LocustID:393303"
/translation="MDSFSTKNAALQAKKMSKMAKTIVANLFIIDTSSEVLDLYR
VTKETRNKRKAOKITIKNLIKRVYKGLVLRNCPNNEELAVERRKKVHTIATNAV
SFYQIDFTFDRRWSMLNDCELIHQAINRHLTAASHARINHVFHFDCCPLATLY
GPSEVYRGHLOKICEGVNKMDEGNL"

ORIGIN
Query Match 42.7%; Score 256.4; DB 5; Length 3986;
Best Local Similarity 66.0%; Pred. No. 2.7e-53;
Matches 371; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 40 CACAGATGCTTTAATTCACAAAACCTGGCCGTTTCAGGACAAAGAAGATCTTGGGTAA 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 CATGACTCGTTGAGCACCAGAACTAGCCCTGAGGGCGCAGAGAAGCTCATGAGCAA 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 AATGATGTCCTCAATTCATGCCACACCTTAATAGACACAAAGTAGAGGTGCTGGA 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 255 GATGGGACCAACACAGTGCCCAACCTTTATAGACACACACAGACGCGAGTACTGGA 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 TGAGCTCTACAGATGACGAGGAGTACACCCAAACAAAGAAGGCGAGAGAAGTCA 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 CGAGCTGTACAGATGACCAAGAGTACAGCGGCAACCGCAAGAGGCCAGAAAGTCA 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 CAAAGACCTTCATCAAGACAGTCAACAGCTGGCATTCTTTATAGAAATATGATTTAA 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 CAAAACTTCATCAAGATGAGTGTCAAGTTGGGGGTCTCTACCGCAAGGTCAATTTCA 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 280 TCAAGATGAGCTTGATTTGAGAAATTTAAGAAAGATTGATCATGCTTCTGATGAC 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 435 CAACGAGAGCTAGCGTTAGTCAAGCGTTTCTTAAGAAAGTGCATACGCTGCGATGAC 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 340 CGTGTCAGTTTCATCAGGTGATTAATACCTTTGACCGGAATGTGTTTACGAGCTGTT 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 AGCTGTAGCTTTACCAATGCACTTCACTTTTGTGACGCGGTCAATGATTAATCTACT 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 AATGATGACAGAGATGCTGACCAAAATCATTTAGCGCCACCTCATGCGCAAGTACA 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 555 CAAATGACTGCGGGAATCTTGACACGAGCATCAATCGGCAATCTAAGCGGAATCTCA 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 TGAAGCGGTAATATATGCTTTGATCAATTTTTCAGATTTGAAATTTTGGCTGCTTGA 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 615 CGCCCAATCAACCAATGCTTCAATCAATTCGCGGATTTGACTTCTCCGACGCTATA 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 TAATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATGATTCACAAAT 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 675 CGGACTTCAGAAAGTAGACCGCGGCACTTGACAGAAATCTGTGAAGAGTCAACAGAT 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 GTTGATGAGAGAAATATGCA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 735 GCTGAGCAAGAGCAATTTTGA 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 34
BX649252/205949 bp DNA linear VRT 06-FEB-2004
LOCUS zebrafish DNA sequence from clone DKEY-49M19 in linkage group 22,
DEFINITION complete sequence.
ACCESSION BX649252
VERSION BX649252.7 GI:42454788
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)

```

```

ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 205949)
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (06-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 6, 2004 this sequence version replaced gi:41392451.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/c/elegans/wormpep/Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhiyong Bao and Sean Eddy, submitted), and those
beginning 'dr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-49M19
is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5.

FEATURES
source
1..205949
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_lib="DKEY-49M19"
/clone_lib="Daniokey"

ORIGIN
Query Match 42.7%; Score 256.4; DB 5; Length 205949;
Best Local Similarity 66.0%; Pred. No. 3e-53;
Matches 371; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 40 CACAGATGCTTTAATTCACAAAACCTGGCCGTTTCAGGACAAAGAAGATCTTGGGTAA 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69527 CATGACTCGTTGAGCACCAGAACTAGCCCTGAGGGCGCAGAGAAGCTCATGAGCAA 69468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 AATGATGTCCTCAATTCATGCCACACCTTAATAGACACAAAGTAGAGGTGCTGGA 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69467 GATGGGACCAACAGAGTGGCCAACTTTTATAGACACACAGACGAGAGTACTGGA 69408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 TGAGCTCTACAGATGACGAGGAGTACCCCAAAACAAAGAAGGAGGCAAGATGAT 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69407 CGAGCTGTACAGATGACCAAGAGTACCGGCAACCGCAAGAGAGGCCAGAAAGTCA 69348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 CAAAGACCTTCATCAAGACAGTCAAGCTGGCATTCTTTATAGAAATATGATTTAA 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 69347 CAAAACCTCATCAAGATGGGTGCTCAAGTTGGGGCTCTCTACCGCAACGGTCAATTCAA 69288
 QY 280 TCAGATGAGCTAGCATTTAGTAGAGAAATTTAGAGAAAGTTCATCAGTTGGTATGAC 339
 Db 69287 CAACGAGAGAGTACGTTAGTCGAGCGTTTTCGTAAAGAGTGCATACGGTGGAGTAC 69228
 QY 340 CTTGCTCACTTTTCATCAGGTGAGTTATTAATCTTTGACCGGAATGTTATCAGGCTGT 399
 Db 69227 AGCTTTAGCTTTCACAGATCGATTCATCTTTGATGAGCGCTCATGAGTAATCTACT 69168
 QY 400 AATGATGACAGAGATGCTGACCAATCATTCAGGCGACCTGCACTGCGCAAGTACA 459
 Db 69167 CAATGACTGCCGTAACTTTCGACACGCGCATCATCGCATTTAACGCGGAATCTCA 69108
 QY 460 TGAACGGGTTAATATGCTTTGATCATTTTTCAGATTTGATGATTTTGGCTGCTTGA 519
 Db 69107 CGCCGAAATCAACATGCTTCATCATTTTCCGCGATTTGATGATTTCCCTCGGAGCTATA 69048
 QY 520 TAATCTTTTGGGAATTTTAAACCCCTTACAAAACCTATGATGATGATCAAAAT 579
 Db 69047 CGACCTTCAGAACTGACCGCGGCACTTGACAGAGATCTGTAGAGAGTCAACAGAT 68988
 QY 580 GTTGATGAGAGAGAACATATGA 601
 Db 68987 GCTGACGAGAGGACATCTTGA 68966

RESULT 35

AC012678 188937 bp DNA linear HTG 18-JUL-2000
 LOCUS AC012678
 DEFINITION Homo sapiens chromosome 15 clone RP11-394B5, WORKING DRAFT
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

AC012678 188937 bp DNA linear HTG 18-JUL-2000
 VERSION AC012678.2 GI:6728953
 KEYWORDS HTG: HTGS PHASE; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE AUTHORS
 Bruno, D., Conn, L., Dela Rosa, M., Federpiel, N., Foreman, P.,
 Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
 Morehouse, A., J., Oefner, P., Palm, C. J., Ramirez, D., Wilhelmy, J.,
 Yu, S., and Davis, R. W.
 Unpublished
 2 (bases 1 to 188937)

JOURNAL REFERENCE
 AUTHORS Bruno, D., Conn, L., Dela Rosa, M., Federpiel, N., Foreman, P.,
 Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
 Morehouse, A., J., Oefner, P., Palm, C. J., Ramirez, D., Wilhelmy, J.,
 Yu, S., and Davis, R. W.
 Direct Submission
 Submitted (03-NOV-1999) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 On Jan 21, 2000 this sequence version replaced gi:6223629.

TITLE

JOURNAL
 COMMENT

COMMENT

Center: Stanford DNA Sequencing and Technology Development
 Center

Center code: SDSTDC
 Web site: http://sequence-www.stanford.edu/group/human/
 Contact: hum-info@sequence.stanford.edu
 Project Information
 Center project name: 698
 Center clone name: RP11-394B5

Summary Statistics
 Sequencing Vector: M13mp18; X02513
 Chemistry: Dye primer; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 177081 bases at least Q40
 Consensus quality: 18532 bases at least Q30
 Consensus quality: 18676 bases at least Q20

Insert size: 185828; agarose-fp
 Insert size: 188487; sum-of-ctnigs
 Quality coverage: 5.3 in Q20 bases; agarose-fp
 Quality coverage: 5.2 in Q20 bases; sum-of-ctnigs.
 NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 ctnigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the ctnigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES

source
 1. 188937
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /clone="RP11-394B5"
 /clone_1ib="RP11 human BAC library 11"

ORIGIN

Query Match 42.6%; Score 255.8; DB 2; Length 188937;
 Best Local Similarity 65.5%; Pred. No. 4,2e-53;
 Matches 374; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

Db 31 TCCGATGCGCAGATGCTTTTATTCAAAACCTGGCGCTCAGGCAAAAAGAT 90
 Db 51681 TACCATAGCTCTGATGTTTTCCTTCAAGAGTCTTGCGCTTCAAGCCCAAGAAAGAT 51740
 QY 91 CTTGGGTAAATGATGTCATCAATCCATCGCCACACCTTAAATAGACGACAAAGTAGTA 150
 Db 51741 TCTGACCAAAATGCGCAGCAAACTGTGGCAACATGTTATGATGATGACACGACGCA 51800
 QY 151 GGTCTGATGAGCTTCAAGAGTACCCAGGAGTACACCCAAACAAAGAGAGGAGA 210
 Db 51801 GATCTTGAATGAGCTCTACAAAGTACCAAGAGCACACACAAAGAGAGGCCA 51860
 QY 211 GAAGATCATCAAGAACCTCATCAAGACAGCAAGCAAGCGCATCTTTATAGAAATA 270
 Db 51861 CAATGATCAAGAAAGACTTATCAAGGTGGGATCAAAATCGGATCCTTACCGAACA 51920
 QY 271 TCAGTTTAAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330
 Db 51921 CAGTTTAAAGCAAGAGAGAGTGTATTTGTGGAGAAATTCGGAAGAGTGAACAGAC 51980
 QY 331 TGTATGACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390
 Db 51981 CGCATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 52040
 QY 391 CAGGCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
 Db 52041 CAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 52100

Qy	451	CAACATCAGCAGGAGGATTAATGATCTTGATCAATTTTCAGATTTGAAATTTTGGC	510
Db	52101	CAGAGACCAACGGGGGCATCAACACACGCTTTTAACACACTTTCGCATGTGGAGTTCTCTC	5216
Qy	511	TGCCTTGATATATCCTTTTGGAAATTTTAAACCCCACTTAACAAAATATGATGATAT	570
Db	52161	CACCCCTATATAGTCGTGATGAGACTGTAGAGCCCAACCTCAAGAGAAATTTGTGAAGAAAT	52222
Qy	571	CAACAAATATGATGATGAAGAAATATGCA	601
Db	52221	CAATPATGTTGCTAGATGAGAAAGTCTTTTAA	52251
RESULT 36			
LOCUS	AC073964/c		
DEFINITION	Homo sapiens chromosome 15 clone RP11-394B5 map 15q21.2, complete sequence.		
ACCESSION	AC073964		
VERSION	AC073964.3	GI:13112225	
KEYWORDS	HTC		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 189796) Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Dots, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A., Madan, A., Nesbitt, R., Traicoff, R. and Hood, L.		
TITLE	Sequencing of human chromosome 15 D15S146-D15S117 region		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 189796) Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradarani, L., Birditt, B., Bloom, S., Dots, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A., Madan, A., Nesbitt, R., Shaffer, T. and Hood, L.		
AUTHORS	Direct Submission		
TITLE	Submitted (08-JUL-2000) Multimegabase Sequencing Center, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105, USA		
JOURNAL			
COMMENT	On Feb 23, 2001 this sequence version replaced gi:9956002.		
----- Genome Center			
Center: Multimegabase Sequencing Center			
Center code: UMWSC			
Web site: http://chroma.mbl.washington.edu/msg_www			
Contact: leerowen@systembiology.org			
Drafting center: SDBTDC			
----- Summary Statistics			
Sequencing vector: pUC18; L08752			
Chemistry: Dye-terminator Big Dye; 90% of reads			
Chemistry: Dye-Primer Big Dye; 10% of reads			
Assembly program: Phrap; version 0.990399			
Note: Date from overlapping BACs AC022407 (drafting center: UMWSC)			
and AC012169 (drafting center: UMWSC) were added for finishing			

Location/Qualifiers			
1. 189796			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
/chromosome="15"			
/map="15q21.2"			
/clone="RP11-394B5"			
/clone.lib="RP11 human BAC library 11"			
/note="Data from overlapping BACs RP11-105D1 and			

[illegible]

	Query Match	42.6%	Score 255.8	DB 9	Length 189796
	Best Local Similarity	65.5%	Pred. No. 4.2e-53		
	Matches 374	Conservative 0	Mismatches 197	Indels 0	Gaps 0
Qy	31	TCGCATGGCCACAGATGCTTTTAATTCAAAAACCTGGCCGTTTCAGGCACAAAAAAGAT	90		
Db	68392	TACCATATGCTCGATGATGTTTATTTAGTTCAAGAGCTTGCCGCTTCAAGCCAGAAAGAT	6833		
Qy	91	CTTGGTAAAAATGGTGTCCAAATTCATCGCCACCCTTAATATGACGACAAAGTAGTA	150		
Db	68332	TCGAGCAAAATATGCGCAGCAAAACCTGGGCAACATGTTATGATATACACGACACGA	6827		
Qy	151	GGTGTGGATGAGTGCTACAGAGTGCACAGAGTACACCCAAACAAAGAGGAGGAGA	210		
Db	68272	GATCTTTGAATGAGCTCTACAAAGTACCAAGACCAAGAGCAACACAAAGAGGAGCCCA	6821		
Qy	211	GAAGATCATCAAGAACCTCATCAAGACATCATCAAGCTGGCACTTTATATAGAAATA	270		
Db	68212	CAAGATCATGAAAGACTTAATCAAGGTGGGAGATCAAAATCGGATCCTTACCGAACA	6815		
Qy	271	TCAGTTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCAACGCT	330		
Db	68152	CAAGTTTAGCCAAAGAGAGCTGGTTATTTGAGGAAGTTCCGGAAGAGCTGAACGAGAC	6809		
Qy	331	TGCTATGACCGGTGTCACTTTCCATCAGGTGATTAATCTTTGACCGGAATGTATTATC	390		
Db	68092	CGCATGACCATTTGTCACTTCTATGAGGTGGAATACACCTTGATGGAACGGTCTCTC	6803		
Qy	391	CAGGCTGTTAAATGAATGACAGAGAGTGTGCACCAATCATTCAGGCGCACCTCACTGC	450		
Db	68032	CAATCTCTCAATGATGACAGAGCTGTGTGCATGAATCTGTGACGGCGACCTGAGGCC	6797		
Qy	451	CAAGTCATGACGACGGGTAAATATGCTTTGATCATTTTTCAAGATGTGAATTTTGGC	510		
Db	67972	CAGAGCCACGAGGCGATCAACACGCTTTTAACAATTTGGCATGTGAAGTTCCCTCTC	6791		
Qy	511	TGCTTGTATATTCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGTAT	570		
Db	67912	CACCTCTATATATGTGATGAGAGCTGTAGGCCCAACCTCAAGAGATTTGTGAAGGAAT	6785		
Qy	571	CAACAAATATGTTGATGAAGAGAACATATGA	601		
Db	67852	CAATTAAGTGTCTAGATGAGAAAGCTTTTAA	67822		
RESULT 37					
LOCUS	CQ736656	1248 bp	DNA	linear	PAT 03-FEB-2004
DEFINITION	Sequence 22590 from Patent WO02068579.				
ACCESSION	CQ736656				
VERSION	CQ736656.1	GI:42332642			
KEYWORDS					

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kites, such as nucleic acid arrays, comprising a majority of humamexons or transcripts, for detecting expression and other uses thereof

JOURNAL Patent: WO 0206579-A 22590 06-SEP-2002;
 PE Corporation (NY) (US)
 Location/Qualifiers
 1..1248
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 42.3%; Score 254.4; DB 6; Length 1248;
 Best Local Similarity 65.9%; Pred. No. 8.2e-53;
 Matches 369; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

42 CAGATGCTTTAATTCCTGCGCCGCTTCAGGCAAAAGAGATCTGGGTAAA 101
 |||||
 689 CTGATGTTTTAGTTCAAGAGCTTGGCGCTTCAAGCCAGAGAAATTTGAGCAAAA 748
 |||||
 102 TGGTGTCAAAATCATTGCGCAACCTTAATAGACGACAGAGTAGTGAGGTGCTGATG 161
 |||||
 749 TAGCCAGCAAAAGCTGTGCGCAACATGTTGATGTATGACACAGCAGCAGAGATCTTTGATG 808
 |||||
 162 AGCTCTACAGAGTACAGGAGGAGTACACCCAAAAGAGAGGAGGAGAGATATCA 221
 |||||
 809 AGCTCTACAAAGTACCAAGAGAGACACACAAAGAGAGAGAGAGAGAGATCA 868
 |||||
 222 AGAAGCTCATCAAGACAGTATGAGTGGCCATTTCTTAAGAGAAATCAAGTTATC 281
 |||||
 869 AAGACTTAATCAAGAGTGGCCATCAAAATCGGATCTCTACCGAACAACCAATTGACC 928
 |||||
 282 AAGATGAGTACATGATGATGAGAGAAATTTAAGAGAAAGTTCATCAGCTTGTATGACG 341
 |||||
 929 AAGAGAGCTGTGTTATGTGAGAGAGTTCGGAAGAGTGAACACAGACCGCATGACCA 988
 |||||
 342 TGGTCAAGTTCCATCAGTGGATTTATACCTTTGACCGGAAATGTGTTATCCAGCTGTAA 401
 |||||
 989 TTGTCAAGCTTTATGAGAGTGAATACCTTCATGAGAAAGTGTCTCTCCATCTCTGTC 1048
 |||||
 402 ATGAATGACAGAGATGCTGACCAAAATCAATTCAGGCGCATCTCATCTGCGCAAGTCACATG 461
 |||||
 1049 ATGAGTGCAGAGACCTGTGTGATGAACTGTGTGACGCGGACCTGACGCCCAAGACCCACG 1108
 |||||
 462 GACGGGTTAAATGCTTTGATCATTCTTCAATGTGTGAATTTTGGCTGCTGTGATA 521
 |||||
 1109 GGGGATCAACACAGCTCTTTAACACCTTTCGCGATGTGAGTTCTCTCCACCTCTATA 1168
 |||||
 522 ATCTTTTGGGAATTTTAAACCCCACTTAACAAAATCATGTGATGTATCAAAAAATGT 581
 |||||
 1169 GTCTGATGAGAGCTGTAGGCCCAACTCAAGAGGATTTGTGAAGATCAATTAAGTTGC 1228
 |||||
 582 TGGATGAAGAGAACATATGA 601
 |||||
 1229 TAGATGAGAAAGTCTTTAA 1248
 |||||

RESULT 38
 CO841622 1986 bp DNA linear PAT 02-AUG-2004
 LOCUS CO841622
 DEFINITION Sequence 269 from Patent EP1440981.
 ACCESSION CO841622.1 GI:50893409
 VERSION CO841622.1
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.
 TITLE Full-length human cdna
 JOURNAL Patent: EP 1440981-A 269 28-JUL-2004;
 Research Association for Biotechnology (JP)
 Location/Qualifiers
 1..1986
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 42.3%; Score 254.4; DB 6; Length 1986;
 Best Local Similarity 65.9%; Pred. No. 8.3e-53;
 Matches 369; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

42 CAGATGCTTTAATTCCTGCGCCGCTTCAGGCAAAAGAGATCTGGGTAAA 101
 |||||
 150 CTGATGTTTTAGTTCAAGAGCTTGGCGCTTCAAGCCAGAGAAATTTGAGCAAAA 209
 |||||
 102 TGGTGTCAAAATCATTGCGCAACCTTAATAGACGACAGAGTAGTGAGGTGCTGATG 161
 |||||
 210 TAGCCAGCAAAAGCTGTGCGCAACATGTTGATGTATGACACAGCAGCAGAGATCTTTGATG 269
 |||||
 162 AGCTCTACAGAGTACAGGAGGAGTACACCCAAAAGAGAGGAGGAGAGATATCA 221
 |||||
 270 AGCTCTACAAAGTACCAAGAGAGACACACAAAGAGAGAGAGAGAGAGATCA 329
 |||||
 222 AGAAGCTCATCAAGACAGTATGAGTGGCCATTTCTTAAGAGAAATCAAGTTATC 281
 |||||
 330 AAGACTTAATCAAGAGTGGCCATCAAAATCGGATCTCTACCGAACAACCAATTGAGCC 389
 |||||
 282 AAGATGAGTACATGATGAGAGAAATTTAAGAGAAAGTTCATCAGCTTGTATGACG 341
 |||||
 390 AAGAGAGCTGTGTTATGTGAGAGAGTTCGGAAGAGCTGAACACAGACCGCATGACCA 449
 |||||
 342 TGGTCAAGTTCCATCAGTGGATTTATACCTTTGACCGGAAATGTGTTATCCAGCTGTAA 401
 |||||
 450 TTGTCAAGCTTTATGAGAGTGAATACACCTTGTATGAGAAAGTGTCTCTCCATCTCTGTC 509
 |||||
 402 ATGAATGACAGAGATGCTGACCAAAATCAATTCAGGCGCATCTCATCTGCGCAAGTCACATG 461
 |||||
 510 ATGAGTGCAGAGACCTGTGTGATGAACTGTGTGACGCGGACCTGACGCCCAAGACCCACG 569
 |||||
 462 GACGGGTTAAATGCTTTGATCATTCTTCAATGTGTGAATTTTGGCTGCTGTGATA 521
 |||||
 570 GGGGATCAACACAGCTCTTTAACACCTTTCGCGATGTGAGTTCTCTCCACCTCTATA 629
 |||||
 522 ATCTTTTGGGAATTTTAAACCCCACTTAACAAAATCATGTGATGTATCAAAAAATGT 581
 |||||
 630 GTCTGATGAGAGCTGTAGGCCCAACTCAAGAGGATTTGTGAAGATCAATTAAGTTGC 689
 |||||
 582 TGGATGAAGAGAACATATGA 601
 |||||
 690 TAGATGAGAAAGTCTTTAA 709
 |||||

RESULT 39
 AK123281 1986 bp mRNA linear PRI 09-SEP-2003
 LOCUS AK123281
 DEFINITION Homo sapiens cDNA FLJ41287 f1s, clone BRAMY2038484, moderately similar to Homo sapiens MDC-3.13 isoform 2 mRNA.
 ACCESSION AK123281
 VERSION AK123281.1 GI:34528784
 KEYWORDS oligo capping, f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saico, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Megatsuma, M., Murakawa, K., Kanehori, K., Takahashi, Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuh, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1986)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - 3' end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1..1986
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRAMT2038484"
 /issue_type="amygdala"
 /clone_lib="BRAMT2"
 /note="cloning vector: pME18SFL3"
 95..709
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC85572.1"
 /db_xref="GI:34528785"
 /translation="MDSDEGESEGEPTAAGDPVPSKSLAQAOKILSKIAKTVAMLLDSTSEIFDELKVKTHKHEHAKHMKMLIYALIGLYVNNPOSOELVVEKRRKLTNOTAMTIVSPYEVYEDPNVLSNLIHECKDYLHELVNHLPRTHGRINHVRHNPADVEFLSTLISLDGDCRNLRICEGINKLDEKYL"
 ORIGIN
 Query Match 42.3%; Score 254.4; DB 9; Length 1986;
 Best Local Similarity 65.9%; Pred. No. 8.3e-53;
 Matches 369; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
 42 CAGATGCTTTTAATTCACAAAACCTGCGCGTTCAGGCAACAAAGATCTGGTAA 101
 150 CTGATGTTTAAAGTCAAGAGCTTGCCTTCAGCCCGAGAGAGATTCGACAAA 209
 102 TGGTTCCAATTCATCGCCACCACTTAATAGACACCAATAGTAGGCTGGATG 161
 210 TAGCCAGCAAACTGTGGCCACATGTATGATATACACACGACGAGATCTTGA 269
 162 AGCTTACAGAGTACAGGAGTACACCAACAAAGAGAGGAGAGAGATCATCA 221
 270 AGCTTACAAAGTACCAAGAGACACACACAAAGAGAGAGAGATCATCA 329
 222 AGAATCTATCAAGACATCAAGCTGCGCATCTTTTATAGAAATCACTTTATC 281
 330 AAGACTTAATCAAGGTGGCGATCAAAATGGGATCTCTTACCGAACAACAGTT 389
 282 AAGATGAGTAGATGATGAGAAATTTAAGAAAGTATCATAGCTTGTATGACCG 341
 390 AAGAGAGAGTGTATTTGAGAAAGTCCGAAAGAGTGAACGACGACCATATCA 449
 342 TGGTCAATTCATCAGTGTGATTAATCTTGAACCGAATGTATCCAGGCTTTAA 401
 450 TTGTGAGCTTCTATAGGTGATATACCTTGATAGGAACGTCTCTCAATCTCTG 509
 402 ATGATGACAGAGATGCTGACCAATCATTCAGGCGCACTGACGCAAGTCAATG 461
 510 ATGAGTCAAGGACCTGTGATGAATCTGATGACGCGGACCTGACGACGACCA 569

462 GACGGTTAATATGCTTTGATCATTTTTCAGATTGGAATTTTGGCTGCTGATA 521
 570 GCGGATACACACAGCTTTTACCACTTTCGCGATGAGATCTCTCCACCTCTATA 629
 522 ATCTTTTGGGAAATTTTAAACCCCACTTACAAAACATATGATGATATCAAAAATG 581
 630 GTCTGATGAGACAGTGTAGGCGCAACCTCAAGAGATTTGTGAAGATCAATAAGTTC 689
 582 TGGATGAGAGAACTATGA 601
 690 TAGATGAGAAAGTCTTTAA 709
 Db
 QY
 582 TGGATGAGAGAACTATGA 601
 690 TAGATGAGAAAGTCTTTAA 709
 RESULT 40
 AJ720336
 LOCUS
 DEFINITION
 Gallus gallus mRNA for hypothetical protein, clone 15114.
 ACCESSION
 AJ720336
 VERSION
 AJ720336.1 GI:53133331
 KEYWORDS
 ORF1.
 SOURCE
 Gallus gallus (chicken)
 ORGANISM
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 REFERENCE
 1
 Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezubov, Y., Zaim, J., Pledier, P., Kutter, S., Blagodatki, A., Kostovska, D., Koter, M., Plachy, J., Carninci, P., Hayashizaki, Y. and Buerstedde, J.M.
 Full-length cDNAs from bursal lymphocytes to facilitate gene function analysis
 Unpublished
 2 (bases 1 to 1589)
 REFERENCE
 Caldwell, R.B.
 Direct Submission
 Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum, Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr. 1, D-85764 Neuherberg, GERMANY
 Location/Qualifiers
 1..1589
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="CB"
 /db_xref="taxon:9031"
 /clone="15114"
 /cell_type="lymphocyte"
 /issue_type="bursa"
 /clone_lib="rikeni"
 /dev_stage="2 weeks old"
 1..1589
 /locus_tag="RCJMB04_15114"
 <1..140
 /locus_tag="RCJMB04_15114"
 141..701
 /locus_tag="RCJMB04_15114"
 /note="ORF1"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAG31995.1"
 /db_xref="GI:53133332"
 /translation="MDTSTKRLALQAKLISKASKTVANVFIIDTSSBLDELTYRATKETHNRKPKOKIKLKIYMLGLVYNGGSPPELTVMBPRKKVHTLMTAVSFHQIDFTFDRRVSSVLTCECDLHQAIVSSHLTKSRIWVNNHNPADVEFLSALYGPADYRTHLKEICGVNKMLEDNI"
 702..>1589
 /locus_tag="RCJMB04_15114"
 ORIGIN
 Query Match 42.0%; Score 252.4; DB 5; Length 1589;
 Best Local Similarity 65.8%; Pred. No. 2.6e-52;
 Matches 367; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
 3'UTR

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sander, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soosa, D., Stead, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Vals, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, D., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 256285)
Worley, K.C.

Direct Submission
Submitted (20-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 256285)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 23, 2002 this sequence version replaced gi:23269911.

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSWY
Center clone name: CH230-77C4
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 199255 bases at least Q40
Consensus quality: 203424 bases at least Q30
Consensus quality: 206379 bases at least Q20
Estimated insert size: 195822; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 24557: contig of 24557 bp in length
* 24558 24657: gap of unknown length
* 24658 201810: contig of 177153 bp in length
* 201811 201910: gap of unknown length
* 201911 214313: contig of 12403 bp in length

FEATURES
source
1..256285
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-77C4"

misc_feature
1..1142
/note="wgs contig"

misc_feature
4086..6215
/note="wgs contig"

misc_feature
24636..25854
/note="wgs contig"

misc_feature
228541..229886
/note="wgs contig"

misc_feature
245270..247000
/note="wgs contig"

ORIGIN
Query Match 41.6%; Score 250.2; DB 2; Length 256285;
Best Local Similarity 64.7%; Pred. No. 1.1e-51;
Matches 312; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

27 CTCCTCCAGTGGCCACAGATGCTTTAATTCACAAAACCTGGCGTTCCAGCACAAGA 86
|||
120612 CTTTACCCATAGATGCTCTCATGTGTAGTCCAAAGTCTCCCTTCAAGCCAGAA 120553
|||
87 AGATCTGGGTAAATNGTGTCAATCATGCGCCACCACTTAATPAGCGACACAAGTA 146
|||
120552 AGATCTGGGTAAATNGTGTCAATCATGCGCCACCACTTAATPAGCGACACAAGTA 120493
|||
147 GTGAGGTGCTGATGATGCTCTACAGATGACCCAGGAGTACACCCAAAAGAAAGAG 206
|||
120492 GCGAATCTTTATATAGCTGTACAAAGTACCCGAAAGAACACCAAGAAAGAG 120433
|||
207 CAGAGAAATCATCAAGAACTCTATCAGACATCATCAAGTGGCCATTCTTTATAGA 266
|||
120432 CCCACAAATCATCAAGAACTCTATCAGACATCATCAAGTGGCCATTCTTTATAGA 120373
|||
267 ATATCAGTTTAATCAAGATGAGTACATTTAGTGAAGAAATTTAAGAAAGTTCAATC 326
|||
120372 ACAAACAGTTCACTCAAGAGAGGTTGTGATTTGAGAAATCTCCGAAAGAACTGAACC 120313
|||
327 AGCTTGCTATGACCGTGTCACTTTCATCAGGTGATTTATACCTTTGACCGAAATGTGT 386
|||
120312 AAGACGCGCATGACATGCTCACTTCTACAGAGTGAATATCCTTTGATAGAAATGTGC 120253
|||
387 TATTCAGGCTGTTAAATGAATCAGAGATGCTGACCAATTCATTACGCGCACCTCA 446
|||
120252 TCTCTAACTTCTGATGATGATCAAGGACCTGTCCATGACCTGACGCACTTGA 120193
|||
447 CTGCAAGTCATGAGACGGGTTAATATGCTTTGATATTTTCAGATTGGAATTTT 506
|||
120192 CGCCCAAGAACCAAGGACGATTAACATGCTTCAACCACTTCGATGTGGAATTTCC 120133
|||
507 TGGCTGCTTGTATATCTTTTGGAAATTTAAACCCCACTTACAAAACATATGATG 566
|||
120132 TTTTCACTTCTACAGTCCGACAGAAATCGAGGCCCAATCGAAGAGATTGGAG 120073
|||
567 GTATCAACAAATGTTGATGAGAAAGAAATATATGA 601
|||
120072 GAATCAATTAATTTGATGATGACAAATCTCTGA 120038
|||

RESULT 43	AC129440	273225 bp	DNA	linear	HTG 13-MAY-2003
LOCUS	AC129440				
DEFINITION	Rattus norvegicus clone CH230-112K14, *** SEQUENCING IN PROGRESS				
ACCESSION	AC129440				
VERSION	AC129440.3				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 273225)				
AUTHORS	Mueny D, Marie, Metzker, M, Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anulepci, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bialwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedetich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dim, H., Diya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duvall, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Georgegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Haravey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, S., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivel, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowls, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheewa, L., Louiseged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mehiney, S., McLeod, M. P., McKell, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokoilehem, O., Okwuonu, G., Olamundugon, A., Pal, S., Parke, K., Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regler, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvatsbeyn, A., Sisson, I., Sitter, C. D., Snaj, D., Sneed, A., Sodegren, E., Song, X. Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svalte, A., Tabot, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., Wiley, F., Williams, G., Willson, R., Wlecyk, R., Wooden, H., Worthy, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.				
TITLE	Unpublished				
JOURNAL	2 (bases 1 to 273225)				
REFERENCE	Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				

```

REFERENCE
TITLE    Rat Genome Sequencing Consortium.
AUTHORS  3 (bases 1 to 273225)
JOURNAL  Submitted (13-May-2003) Human Genome Sequencing Center, Department
        of Molecular and Human Genetics, Baylor College of Medicine, One
        Baylor Plaza, Houston, TX 77030, USA
COMMENT  The sequence in this assembly is a combination of BAC based reads
        and whole genome shotgun sequencing reads assembled using Atlas
        (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
        in the feature table below represents a scaffold in the Atlas
        assembly (a 'contig-scaffold'). Within each contig-scaffold,
        individual sequence contigs are ordered and oriented, and separated
        by sized gaps filled with Ns to the estimated size. The sequence
        may extend beyond the ends of the clone and there may be sequence
        contigs within a contig-scaffold that consist entirely of whole
        genome shotgun sequence reads. Both end sequences and whole genome
        shotgun sequence only contigs will be indicated in the feature
        table.

        ----- Genome Center
        Center: Baylor College of Medicine
        Center code: BCM
        Web site: http://www.hgsc.bcm.tmc.edu/
        Contact: hgsc-help@bcm.tmc.edu

        ----- Project Information
        Center project name: GKTY
        Center clone name: CH230-112K14

        ----- Summary Statistics
        Assembly program: Atlas 3.0:
        Consensus quality: 224959 bases at least Q40
        Consensus quality: 227786 bases at least Q30
        Consensus quality: 229294 bases at least Q20
        Estimated insert size: 236137; sum-of-contigs estimation
        Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

        -----
        * NOTE: Estimated insert size may differ from sequence length
        * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 7 contigs. The true order of the pieces
        * is not known and their order in this sequence record is
        * arbitrary. Gaps between the contigs are represented as
        * runs of N, but the exact sizes of the gaps are unknown.
        * This record will be updated with the finished sequence
        * as soon as it is available and the accession number will
        * be preserved.

        1. 4842: contig of 4842 bp in length
        * 4843 4942: gap of unknown length
        * 4943 265126: contig of 260184 bp in length
        * 265127 265226: gap of unknown length
        * 265227 266729: contig of 1503 bp in length
        * 266730 266829: gap of unknown length
        * 266830 268969: contig of 2040 bp in length
        * 268970 270127: gap of unknown length
        * 270128 270227: gap of unknown length
        * 270228 271518: contig of 1291 bp in length
        * 271519 271618: gap of unknown length
        * 271619 273225: contig of 1607 bp in length.

        Location/Qualifiers
        1. 273225
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-112K14"
        complement(4198..4940)
        /note="clone_boundary"
        clone_end:17
        site:Scori
        end_sequence:BH258362"
        4943 46894
        /note="wgs_contig"

```


	Query Match	41.6%	Score 250.2;	DB 2;	Length 273225;	
	Best Local Similarity	64.7%;	Pred. No. 1.le-51;			
	Matches 372;	Conservative	0;	Mismatches 203;	Indels 0;	Gaps 0;
OY	27 CTCCCTCCAGTGGGCCACGATGTCTTTAATTCCCAAAACCCTGGCCCTTCAGGCACAAGA 86					
Db	8649 CTTTAGCATTAAGTGCTCCTCAATGTTTTACTCCAAAGTCGTGCGCTTCAAAGCCAGAAG 8708					
OY	87 AGATCTTGAGTAAAAATGGTGTCCAAATTCATGCCCCAACCTTTAATAAGCACACAGTA 146					
Db	8709 AGATCTTGAGCAAGATATACCCAGCAAAACCTGGCCCAATGCTGATGATGACACACAGA 8768					
OY	147 GTGAGTGCTGGANAGCTTACAGAGTGACACAGGAATACACCCAAAAGAAGAGNG 206					
Db	8769 GCGAATCTTTGATGATAGCTGTACAAAGTCCACCGAAGAACACACCACAAGAAAGAGNG 8828					
OY	207 CAGAGAAATCATCAAGAACCCTCATCAAGACAGTCATCAAGCTGGSCAYTCTTTAATAGA 266					
Db	8829 CCCHACAAATCATGAAGATGTMAATMAAGTGGCATCAAAAATTGTAITTTCTTACCGGA 8888					
OY	267 ATATACATTTAATCAAGATGAGTAGCATTTGATGAGAGAAATTTAAGAAGAAAGTTCA TC 326					
Db	8889 ACAAACATTCAGTCAMAGAGAGGTGTGATGTGAGAAATCTCGAAGAAATCGAAC 8948					
OY	327 AGCTTGCTATGACCGTGTGCTCAGTTTCCAATCAGTGGATTTATACCTTTGACCGGAATGTGT 386					
Db	8949 AGACCGCATGACAAVGTGTGAGCTTCTACGAAGTGGAGTATACCTTTATACGAATGTGC 9008					
OY	387 TATCCAGGCTGTAAATGATGACAGAGAGATGCTGCAACCAATCAATTCAGCGCCACTCA 446					
Db	9009 TTCTTAACTTCTGCATGATGATGACAGAGACCTGTGTCAATGATGATGATGATGAT 9068					
OY	447 CTGCCAAGTCACATGAGACGGGTAAATATGTCTTTGATCATTTTTCAGATTGTGAATTT 506					
Db	9069 CGCCAGAACCCACGAGCGCATTAACCAATGTCCTTCAACCATCTGCTGATGTGAATTC 9128					
OY	507 TGCTGCTGTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATG 566					
Db	9129 TTTTCACCTCTCTACAGTCCGACGGAACCTGACGGCCCAATCTCAAGAGATTTGCCAAG 9188					
OY	567 GTATCAACAAATGTTGGATGAGAGAACATATGA 601					
Db	9189 GAATCAATTAATTTGTAATGATGCAAAATCTCTTGA 9223					
RESULT 44						
BC076797		1938 bp	mRNA	linear	VRT 20-SEP-2004	
LOCUS	Xenopus laevis MCC83729 protein, mRNA (CDNA clone MGC:83729					
DEFINITION	IMAGE:633647), complete cds.					
ACCESSION	BC076797					
VERSION	BC076797.1 GI:49899742					
KEYWORDS	MGC.					
SOURCE	Xenopus laevis (African clawed frog)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Xenopodinae; Xenopus; Xenopus. 1 (bases 1 to 1938) Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P. Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative Dev. Dyn. 225 (4), 384-391 (2002)					
TITLE	JOURNAL PUBMED REFERENCE AUTHORS					
	2 (bases 1 to 1938) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Hoskins,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schneitz,T.E., Brownstein,M.J., Ueda,I.T.B., Yoshizuki,S.,					

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT
 Email: cga@bs-remail.nih.gov
 Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcag.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
 Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Telka Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu,
 Parvaneh Saedi, JR Santos, Angelique Scherch, Ursula Skalska,
 Duane Smalins, Jeff Stolt, Miranda Tsai, George Yang, Jacque
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAK Plate: 156 Row: d Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.
 Location/Qualifiers
 1. 1938
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="MGC:83729 IMAGE:6639647"
 /tissue_type="Oocytes"
 /clone_1kb="NICHD_XGC_OO1"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6.1"
 1. 1938
 /gene="MGC83729"
 /db_xref="LOCUSID:446389"
 94. 702
 /gene="MGC83729"
 /codon_start=1
 /product="MGC83729 protein"
 /protein_id="AAH76797.1"
 /db_xref="GI:49899743"
 /db_xref="LOCUSID:446389"
 /translation="MDDTSDGDSBGLSGPSPPEPSSKSPVAVQAKTILSMATYTMAN
 MLIDDTSEIIFDELKVTYKEVYVKNKSAHNVLDLVAVAKVGLYVNNKQPSSELEI
 VENNRRKINQTCMTAVSPFVEVYFDGNVSLGLHECKTLLHGLVQHLTPKSHSRID
 RVFHFADVEFLTALVSLBGNVRYRLKICEGVNKLDDERVL"
 38.5%; Score 231.2; DB 5; Length 1938;

Best Local Similarity 64.3%; Pred. No. 5,3e-47;
Matches 347; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

```

QY 50 TTTAATTCAAAAAAGCTGGCGCTTCAGGCAAAAGAAAGTCTTGAGTAAATGATGTC 109
    |||
Db 151 TTTAGCTCAAAAGAGCTTGTCTGTTTCAAGCCAGAGAAAGATTCTCAGTAGATGGCAACA 210
    |||
QY 110 AAATTCATGGCCACCACTTAAATAGAGACAAAGTAGTGTGAGTGTGATGAGCTTAC 169
    |||
Db 211 AAACCAATGGCCCAACATGCTTATGATGACAAAGCAAGTAAATATTGATGAGCTCTTT 270
    |||
QY 170 AGAGTGCACGAGGAGTAGTACACCAAAACAGAGAGGAGAGATCATCAAGAACTTC 229
    |||
Db 271 AAAGTAACCAAGAAATATGTAAAGAACAGAAAGAGGCCCAAGGCTCTGAAAGACCTT 330
    |||
QY 230 ATCAAGACAGTCAATCAAGCTGCGCATCTTTATAGAAATATCAATTAATCAAGATGAG 289
    |||
Db 331 GTAAAGGTTGACAGTAAAGTGGCATTTCTGACGAATTAACAGTTCACTGGAAGAA 390
    |||
QY 290 CTAGCATGATGAGAAATTTAAGAGAAAGTTCAATCAGCTTGCTATGACCGTGTGAGT 349
    |||
Db 391 CTGAATATGCTCAAAAACCTCAGAGAAAGAACTGAATCAGACTGACATGCAAGTCAGT 450
    |||
QY 350 TTCATCAGGTGATTAATACCTTGAACCGGAATGTTATCCAGGCTGTAATTAATGATGC 409
    |||
Db 451 TTTCTTGAAGTTAGTACAGTGTGATTAAGTAAAGTGTCTTCCGACGTTGACAGATGT 510
    |||
QY 410 AGAGAGATGCTGACCAAAATCATTCAGCGCCACCTCCTCAAGTACATGAGAGGAGT 469
    |||
Db 511 AAACCTCTTCTTCAATGATTTGTGACGCGCATCTAACAACCAATCCCACTGCGATC 570
    |||
QY 470 AATTAATGCTTTGATCATCTTTTTCAGATTGGAATTTTGGCTGCTGTAATTCCTTTT 529
    |||
Db 571 GACAGGGTTTATATCATCTTTTGTGATGTGAATTCCTTCTGCTCTATGACCTTGAA 630
    |||
QY 530 GGAATTTTAAACCCCACTTACAAAACATATGATGATGATGATCAAAAGTTGATGAA 589
    |||
Db 631 GGAATTTACAGCGCTGACCTCAAAAAGATTTGTGAAGGGGTCAATTAATTAATGATGAA 690
    |||

```

```

RESULT 45
BC053167 1813 bp mRNA linear VRT 30-JUN-2004
DEFINITION Danio rerio zgc:63960. mRNA (cDNA clone MGC:63960 IMAGE:6791907),
complete cds.
ACCESSION BC053167
VERSION BC053167.1 GI:31418810
KEYWORDS MGC:
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
REFERENCE 1 (bases 1 to 1813)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heish,F.,
Dietzenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneitz,T.E., Brownstein,M.J., Udwin,T.B., Toshlyuk,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettman,M., Madan,A., Kodrzycki,S.,
Sanchez,A.G., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Boutard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length

```

JOURNAL human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1813)
AUTHORS Strausberg,R.
TITLE Direct Substitution
JOURNAL Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Offices, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov

Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

FEATURES
source
1.1813
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:63960 IMAGE:6791907"
/rissue_type="Kidney, zebrafish"
/clone_id="NCI_CGAP_ZK1d1"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6.1"

gene
1.1813
/gene="zgc:63960"
/note="synonym: MGC63960"
/db_xref="LocusID:393322"
178.786
/gene="zgc:63960"
/codon_start=1
/product="hypothetical protein MGC63960"
/protein_id="AAH53167.1"
/db_xref="GI:31418811"
/db_xref="LocusID:393322"
/translation="MDSGQSEBGLSPHESFNSKSLALQAKKILSKMATVAVN
LITDPSSEILBELYASREYTKSKEARKIKDVYKILKIGILVRNHPQFSEMET
VRFKKMNAATVSPFYEVETPDRLGILSTELLCRLHLEVHHLTMSHGRID
HYFNHPADVPLTELGPSEDRILNRKICDGINLIDBGL"

CDS

ORIGIN
Query Match 37.5%; Score 225.4; DB 5; Length 1813;
Best Local Similarity 63.3%; Pred. No. 1.5e-45;
Matches 346; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

QY 44 GATGCTTTAATTCAAAAAAGCTGGCGCTTCAGGCAAAAGAAAGTCTTGAGTAAATG 103
    |||
Db 229 GAGAGCTTCAATTCCAAAGATTGGCCCTTCAGGCTCAAAAGAAAGATTGAGTAAATG 288
    |||
QY 104 GTGTCCAATTCATGGCCCAACCTTAATAGACACACAAGTAGTGAAGTGTGATGAG 163
    |||
Db 289 GCCACATGCGCGTGGCGAAGCTCTTAACAGAGACACAGCGAGATTCGAGCGAA 348
    |||
QY 164 CTCTACAGAGTACCGAGGAGTACACCAAAACAGAGAGGAGCAGAGATCATCAAG 223
    |||
Db 349 CTCTAAGAGCAGTGAATATACCAAGAGAGAGAGCCCAAGATCATCAAA 408
    |||
QY 224 AACCTATCAAGACAGTCAATCAAGCTGCGCATTTCTTATAGAAATATCAAGTTAATCA 283
    |||

```


CC treatment of cancer, tumours, developmental abnormalities and foetal
 CC deficiencies, blood disorders, diseases of the immune system, autoimmune
 CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
 CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
 CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
 CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
 CC also useful for identifying their binding partners. The sequences given
 CC in AAX97907 to AAX97915 and AAY36223 are used in the exemplification of
 CC the present invention

XX Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;

Query Match 97.9%; Score 588.2; DB 2; Length 1943;

Best Local Similarity 99.3%; Pred. No. 1.4e-163;

Matches 590; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GTCCGGGCGCCGTCGCGCACTCTCCGATGGCCACAGATGCTTTTAATCCAAAACCTG 67
 DB 93 GTCCGGGCGCCGTCGCGCACTCTCCGATGGCCACAGATGCTTTTAATCCAAAACCTG 152
 QY 68 GCGGTTCAAGCAAAAGAGATCTGGGTAAATGTGCTCCAAATCCATCCGACACACC 127
 DB 153 GCGGTTCAAGCAAAAGAGATCTGGGTAAATGTGCTCCAAATCCATCCGACACACC 212
 QY 128 TTAATGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGTAC 187
 DB 213 TTAATGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGTAC 272
 QY 188 ACCCAAAACAAAGAGAGGAGGAGAAAGATCATCAAGACCTCTCAAGACAGTCAAG 247
 DB 273 ACCCAAAACAAAGAGAGGAGGAGAAAGATCATCAAGACCTCTCAAGACAGTCAAG 332
 QY 248 CTGGCCATCTTTATGAAATATGATGATTAATCAAGATGAGTACATTTATGAGAAA 307
 DB 333 CTGGCCATCTTTATGAAATATGATGATTAATCAAGATGAGTACATTTATGAGAAA 392
 QY 308 TTTAAGAGAAAGTTGATCATGAGCTTGCTATGACCGGTGATGATTCATGAGTATAT 367
 DB 393 TTTAAGAGAAAGTTGATCATGAGCTTGCTATGACCGGTGATGATTCATGAGTATAT 452
 QY 368 ACCTTGACCGGAAATGTTTATTCAGGCTGTTAAATGACAGAGATGCTGCACAA 427
 DB 453 ACCTTGACCGGAAATGTTTATTCAGGCTGTTAAATGACAGAGATGCTGCACAA 512
 QY 428 ATCATTCAGCGCACTCTCACTGCGCAAGTCAATGACGAGGTTAATATGCTTATCAT 487
 DB 513 ATCATTCAGCGCACTCTCACTGCGCAAGTCAATGACGAGGTTAATATGCTTATCAT 572
 QY 488 TTTTCAGATTTGAAATTTTGGCTGCTGTATATCTTTTGGAAATTTTAAACCCAC 547
 DB 573 TTTTCAGATTTGAAATTTTGGCTGCTGTATATCTTTTGGAAATTTTAAACCCAC 632
 QY 548 TTAACAACCTATGTGATGATGATCAACAAATGTTGATGAGAGAGAACTATGA 601
 DB 633 TTAACAACCTATGTGATGATGATCAACAAATGTTGATGAGAGAGAACTATGA 686

RESULT 4

ADAl1570

ID ADAl1570 standard; DNA; 1943 BP.

AC ADAl1570;

XX 06-NOV-2003 (first entry)

DE Human cDNA encoding a novel secreted protein, SEQ ID NO 98.

XX cancer; inflammation; immune disorder; neurological disorder;

KW blood clotting disorder; food additive; food preservative;

XX storage capability; fat content; nutritional component; ds; gene; human.

XX Homo sapiens.

PN US2003055236-A1.

XX 20-MAR-2003.

XX 14-MAR-2002; 2002US-00097065.

XX 18-DEC-1997; 97US-0068006P.

XX 18-DEC-1997; 97US-0068007P.

XX 18-DEC-1997; 97US-0068008P.

XX 18-DEC-1997; 97US-0068003P.

XX 18-DEC-1997; 97US-0068054P.

XX 18-DEC-1997; 97US-0068057P.

XX 18-DEC-1997; 97US-0068064P.

XX 18-DEC-1997; 97US-0070923P.

XX 19-DEC-1997; 97US-0068169P.

XX 19-DEC-1997; 97US-0068367P.

XX 19-DEC-1997; 97US-0068367P.

XX 19-DEC-1997; 97US-0068369P.

XX 17-DEC-1998; 96WO-US027059.

XX 17-JUN-1998; 99US-00334595.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA, Ruben SM, Carter KC, Shi Y, Rosen CA, Soppet DR,

XX Kyaw H, Wei Y, Florence KA, Duan DR, Florence C, Greene JM, Feng P,

XX Ferlie AM, Yu G, Janat F, Ni J;

XX WPI; 2003-567105/53.

XX P-PSDB; ADAl1694.

XX New secreted HKABT24 nucleic acid molecules and polypeptides, useful for

XX preventing, treating, or ameliorating a medical condition, such as

XX cancer, inflammation, immune disorders, neurological and blood clotting

XX disorders.

XX Claim 1; SEQ ID NO 98; 118pp; English.

XX The invention relates to an isolated HKABT24 nucleic acid molecule. The

XX polypeptides, nucleic acids and antibodies are useful for diagnosing a

XX pathological condition or a susceptibility to a pathological condition,

XX for preventing, treating, or ameliorating a medical condition, such as

XX cancer, inflammation and other immune disorders, neurological and blood

XX clotting disorders. The nucleic acids are also useful for chromosome

XX identification, radiation hybrid mapping or long-range restriction

XX mapping. The polypeptides and antibodies are useful for providing

XX immunological probes for differential identification of the tissues

XX immunohistochemistry assays. The polypeptide, polynucleotide, agonist or

XX antagonist may also be used as a food additive or preservative to

XX increase or decrease storage capabilities, fat content or other

XX nutritional components. The present sequence represents cDNA encoding a

XX novel human secreted protein. Note: The sequence data for this patent did

XX not form part of the printed specification but was obtained in electronic

XX format directly from USPTO at

XX segdata.uspto.gov.uk/sequence.html?DocID=20030055236.

XX Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;

Query Match 97.9%; Score 588.2; DB 9; Length 1943;

Best Local Similarity 99.3%; Pred. No. 1.4e-163;

Matches 590; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GTCCGGGCGCCGTCGCGCACTCTCCGATGGCCACAGATGCTTTTAATCCAAAACCTG 67
 DB 93 GTCCGGGCGCCGTCGCGCACTCTCCGATGGCCACAGATGCTTTTAATCCAAAACCTG 152
 QY 68 GCGGTTCAAGCAAAAGAGATCTGGGTAAATGTGCTCCAAATCCATCCGACACACC 127
 DB 153 GCGGTTCAAGCAAAAGAGATCTGGGTAAATGTGCTCCAAATCCATCCGACACACC 212
 QY 128 TTAATGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGTAC 187
 DB 213 TTAATGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGTAC 272

QY 188 ACCCAAAACAGAGAGGAGAGAGATATCATAGACCTCTCATAGACATCATCAG 247
 DB 273 ACCCAAAACAGAGAGGAGAGAGATATCATAGACCTCTCATAGACATCATCAG 332
 QY 248 CTGGCCATCTTTATAGAGATTAATCACTTTAATCAAGATGAGCATGATGAGAA 307
 DB 333 CTGGCCATCTTTATAGAGATTAATCACTTTAATCAAGATGAGCATGATGAGAA 392
 QY 308 TTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTCCATCAGTGTAT 367
 DB 393 TTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTCCATCAGTGTAT 452
 QY 368 ACCTTTGACCGGAATGTTTATCCAGGCTGTAAATGATGACAGAGATGTCACAA 427
 DB 453 ACCTTTGACCGGAATGTTTATCCAGGCTGTAAATGATGACAGAGATGTCACAA 512
 QY 428 ATCATTCAGCGCACCTCATCTGCAAGTCAATGACAGCGGTTAATATGCTTGTAT 487
 DB 513 ATCATTCAGCGCACCTCATCTGCAAGTCAATGACAGCGGTTAATATGCTTGTAT 572
 QY 488 TTTTCAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGGAATTTAAACCCAC 547
 DB 573 TTTTCAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGGAATTTAAACCCAC 632
 QY 548 TTTCAAAACTATGTGATGTATCAACAAATGTTGGATGAAGAAATATATGA 601
 DB 633 TTTCAAAACTATGTGATGTATCAACAAATGTTGGATGAAGAAATATATGA 686

RESULT 5

ADAS6241
 ID ADAS6241 standard; DNA; 1943 BP.

AC ADAS6241;
 XX

DT 20-NOV-2003 (first entry)
 XX

DE Gene encoding human secreted protein #420.
 XX

XX immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;
 KM cytoskeletal; cerebroprotective; neuroprotective; nootropic;
 KM cardiovascular; antiarteriosclerotic; gene therapy;
 KM human secreted protein; immune disorder; inflammation;
 KM respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
 KM inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
 KM multiple sclerosis; ischemic brain injury; Parkinson's disease;
 KM Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 KM triple helix formation; antisense gene therapy; forensic biology; ds;
 KM gene.

XX Homo sapiens.
 OS

PN MO2002102994-A2.
 XX

PD 27-DEC-2002.
 XX

PF 19-MAR-2002; 2002MO-US008278.
 XX

PR 21-MAR-2001; 2001US-0277340P.
 XX

PR 13-JUL-2001; 2001US-0306171P.
 XX

PR 13-NOV-2001; 2001US-0331287P.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI Rosen CA, Ruben SM;
 XX

DR WPI, 2003-167512/16.
 XX

DR P-PSDB; ADAS7137.
 XX

PT New human secreted polypeptides and polynucleotides, useful for
 PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
 PT conditions, respiratory disorders, cancers, CNS disorders, or

PT neurodegenerative disorders.

XX Claim 21; SEQ ID NO 430; 1754bp; English.

PS The invention relates to 592 new human secreted polypeptides useful for
 XX diagnosing, treating or preventing e.g. immune disorders, inflammatory
 CC conditions, respiratory disorders, cancers, CNS disorders, or
 CC neurodegenerative disorders, or polypeptides comprising an amino acid
 CC sequence at least 95% identical to the new sequences. The polypeptides,
 CC antibodies or antibody fragments that bind to the polypeptides, nucleic
 CC acids encoding the polypeptides, agonists or antagonists that binds to
 CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
 CC compositions for diagnosing, treating or preventing an e.g. immune
 CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
 CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
 CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
 CC (e.g. multiple sclerosis or ischemic brain injury), neurodegenerative
 CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
 CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
 CC polynucleotides are useful for chromosome identification, chromosome
 CC mapping, for controlling gene expression through triple helix formation
 CC or antisense DNA or RNA, in gene therapy, for identifying individuals
 CC from minute biological samples, in forensic biology, and as hybridization
 CC probes. The polypeptides are useful for as molecular weight markers on
 CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
 CC gels, to raise antibodies, for testing biological activities, and for
 CC treating or preventing neural disorders, immune system disorders,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal, proliferative and/or cancerous diseases. This sequence corresponds
 CC to a gene encoding one of the polypeptide of the invention. Note: The
 CC sequence data for this patent did form part of the printed specification,
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1943 BP; 629 A; 325 C; 399 G; 566 T; 0 U; 4 Other;

XX Query Match 97.9%; Score 588.2; DB 10; Length 1943;

XX Best Local Similarity 99.3%; Pred. No. 1.4e-163;

XX Matches 590; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GTCCGGGGCGCGCGCGCACTCCCGATGGCCACAGATGCTTTAATCCAAAACCTG 67
 DB 93 GTCCGGGGCGCGCGCGCACTCCCGATGGCCACAGATGCTTTAATCCAAAACCTG 152
 QY 68 GCGGTTCAGGACCAAAAGAAAGATCTTGGTAAATGCTGCCAATCCATCGGCACACC 127
 DB 153 GCGGTTCAGGACCAAAAGAAAGATCTTGGTAAATGCTGCCAATCCATCGGCACACC 212
 QY 128 TTAATAGACGACCAAGTGTGAGGTGCTGATGAGCTTACAGAGTGACGAGGATAC 187
 DB 213 TTAATAGACGACCAAGTGTGAGGTGCTGATGAGCTTACAGAGTGACGAGGATAC 272
 QY 188 ACCCAAAACAGAGAGGAGGACAGAGATCATCAAGACCTCATCAAGACATCAAG 247
 DB 273 ACCCAAAACAGAGAGGAGGACAGAGATCATCAAGACCTCATCAAGACATCAAG 332
 QY 248 CTGGCCATCTTTATAGAGATTAATCACTTTAATCAAGATGAGCATGATGAGAA 307
 DB 333 CTGGCCATCTTTATAGAGATTAATCACTTTAATCAAGATGAGCATGATGAGAA 392
 QY 308 TTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTCCATCAGTGTAT 367
 DB 393 TTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTCCATCAGTGTAT 452
 QY 368 ACCTTTGACCGGAATGTTTATCCAGGCTGTAAATGATGACAGAGATGTCACAA 427
 DB 453 ACCTTTGACCGGAATGTTTATCCAGGCTGTAAATGATGACAGAGATGTCACAA 512
 QY 428 ATCATTCAGCGCACCTCATCTGCAAGTCAATGACAGCGGTTAATATGCTTGTAT 487
 DB 513 ATCATTCAGCGCACCTCATCTGCAAGTCAATGACAGCGGTTAATATGCTTGTAT 572
 QY 488 TTTTCAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGGAATTTAAACCCAC 547

```
|||||
DB 573 TTTTCAGATGATGTAATTTTGGCTGCTGTGTAATTCCTTTGGGAAATTTTAAACCCAC 632
QY 548 TTAGCAAAACTATGATGATGATCAACAATAATGTTGAGAGAGAACTATGCA 601
DB 633 TTAGCAAAACTATGATGATGATCAACAATAATGTTGAGAGAGAACTATGCA 686

RESULT 6
AAH14888
ID AAH14888 standard; cDNA; 1921 BP.
AC AAH14888;
DT 26-JUN-2001 (first entry)
DE Human cDNA sequence SEQ ID NO:12751.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
PD 07-FEB-2001.
PF 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-MAY-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INSTR.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesising polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX Claim 8; SEQ ID NO 12751; 2537BP + Sequence listing; English.
XX The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX Sequence 1921 BP; 610 A; 334 C; 391 G; 586 T; 0 U; 0 Other;
```

```
Query Match 97.6%; Score 586.8; DB 4; Length 1921;
Best Local Similarity 99.5%; Pred. No. 3.6e-163;
Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CTTGACGCTGCCGGGCGCGTCCG-CGACTCCTCCGATGGCCACAGATGCTTTAATTCGA 59
DB 105 CTTGACGCTGCCGGGCGCGTCCGCGCACTCCCATGCGCCACAGATGCTTTAATTCGA 164
QY 60 AAAACCTGGCGCTTACAGGACAAAAGAGATCTTGGTAAATGGTGTCCAAATCCATCG 119
DB 165 AAAACCTGGCGCTTACAGGACAAAAGAGATCTTGGTAAATGGTGTCCAAATCCATCG 224
QY 120 CCACCACTTAAATAGACGACACAGATAGTGAAGTCTGATGAGCTTACAGAGTACCA 179
DB 225 CCACCACTTAAATAGACGACACAGATAGTGAAGTCTGATGAGCTTACAGAGTACCA 284
QY 180 GGGAGTACACCCAAAACAAAGAGAGGAGAGAGATCATCAAGAACTCATCAAGACAG 239
DB 285 GGGAGTACACCCAAAACAAAGAGAGGAGAGAGATCATCAAGAACTCATCAAGACAG 344
QY 240 TCATCAAGCTGGCCATCTTTTATAGAAATATCAGTTAATCAAGATGAGTAGATTGA 299
DB 345 TCATCAAGCTGGCCATCTTTTATAGAAATATCAGTTAATCAAGATGAGTAGATTGA 404
QY 300 TGGAGAAATTTAAGAAAGTTTCATCAGCTTGTATGACCGTGTGATTCATCAGG 359
DB 405 TGGAGAAATTTAAGAAAGTTTCATCAGCTTGTATGACCGTGTGATTCATCAGG 464
QY 360 TGGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAATGCAAGAGATGC 419
DB 465 TGGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAATGCAAGAGATGC 524
QY 420 TGCACCAATCATTCAGCGCCGACCTCACTGCGCAATGACATGAGCGGTTAATATGCT 479
DB 525 TGCACCAATCATTCAGCGCCGACCTCACTGCGCAATGACATGAGCGGTTAATATGCT 584
QY 480 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTTTA 539
DB 585 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTTTA 644
QY 540 AACCCCATTTACAAAATGATGTGATGATTAACAAAATGTTGATGAGAGAACTAT 599
DB 645 AACCCCATTTACAAAATGATGTGATGATTAACAAAATGTTGATGAGAGAACTAT 704
QY 600 GA 601
DB 705 GA 706

RESULT 7
AAH14793
ID AAH14793 standard; cDNA; 1729 BP.
AC AAH14793;
DT 26-JUN-2001 (first entry)
DE Human cDNA sequence SEQ ID NO:12578.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
PD 07-FEB-2001.
PF 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-MAY-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
```


PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX MPI, 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 8; SEQ ID NO 12578; 2537bp + Sequence listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 SQ Sequence 1729 BP; 563 A; 277 C; 354 G; 535 T; 0 U; 0 Other;
 XX
 XX
 Query Match 93.9%; Score 564.4; DB 4; Length 1729;
 Best Local Similarity 99.8%; Pred. No. 1.5e-156;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 456 CACATGACGGGGTATATATGCTTATGATCATTTTCAATTTGTAATTTGGCTCCCT 515
 DB 522 CACATGACGGGGTATATATGCTTATGATCATTTTCAATTTGTAATTTGGCTCCCT 581
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATCAACA 575
 DB 582 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATCAACA 641
 QY 576 AAATGTTGATGAAGAAGAACATATGA 601
 DB 642 AAATGTTGATGAAGAAGAACATATGA 667
 RESULT 8
 ID ADR14214
 XX ADR14214 standard; DNA; 1814 BP.
 XX
 AC ADR14214;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human NF-kappaB pathway-associated gene SeqID215.
 XX
 KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virologic;
 KW antitubercic; antirheumatic; gastrointestinal-gen; antiaesthetic;
 KW antidiabetic; antihypertensive; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnary; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked amidiotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAB;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; gene; ds; human.
 OS
 XX Homo sapiens.
 XX
 XX
 PN WO2004065577-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004WO-US000798.
 XX
 PR 14-JAN-2003; 2003US-0440068P.
 PR 12-MAY-2003; 2003US-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Nadler SG, Neubauer MG, Feder JM, Carman J;
 XX
 DR MPI; 2004-562168/54.
 DR P-PSDB; ADR14215.
 XX
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 PS Claim 1; SEQ ID NO 215; 237bp; English.
 XX
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virologic, antitubercic, antirheumatic,
 CC gastrointestinal-gen, antiaesthetic, antidiabetic, antihypertensive,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnary activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions

or diseases associated with the NF-kappaB pathway. The condition is an immune disorder, an inflammatory disorder, cancer, aberrant apoptosis, related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas, hematopoietic tumours, hyper-IgM syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell survival, evasion of immune responses, rheumatoid arthritis, inflammatory bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper immune activity, disorders related to aberrant acute phase responses, hypercongenital conditions, birth defects, necrotic lesions, wounds, organ transplant rejection, conditions related to organ transplant rejection, disorders related to aberrant signal transduction, proliferating disorders, cancers and HIV propagation in cells infected with other viruses. The present sequence is that of a human gene which is subject to the novel association with the NF-kappaB pathway of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer from Genbank.

Sequence 1814 BP; 590 A; 297 C; 372 G; 555 T; 0 U; 0 Other;

Query Match 93.9%; Score 564.4; DB 13; Length 1814;
Best Local Similarity 99.8%; Pred. No. 1.6e-156;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 36 TGGCCACAGATGCTTTAAATTCGAAAACCTGGCGGTTCAGGACAAAAGAAATCTTGG 95
DB 175 TGGCCACAGATGCTTTAAATTCGAAAACCTGGCGGTTCAGGACAAAAGAAATCTTGG 234
QY 96 GTAAAAATGATGCTCAATTCATCGCACCACTTAAATAGACACAAAGTAGAGATGC 155
DB 235 GTAAAAATGATGCTCAATTCATCGCACCACTTAAATAGACACAAAGTAGAGATGC 294
QY 156 TGGATAGCTCTACAGAGTAGACCAAGAGTAGACCCAAAACAAGAGAGCAGAGAAG 215
DB 295 TGGACGAGCTCTACAGAGTAGACCAAGAGTAGACCCAAAACAAGAGAGCAGAGAAG 354
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGTGGCGCACTTTTAAAGAAATCAAG 275
DB 355 TCATCAAGAACCTCATCAAGACAGTCATCAAGTGGCGCACTTTTAAAGAAATCAAG 414
QY 276 TTAATCAAGTAGAGTAGAGTAGAGTAAGAAATTTAAGAAAGATTCATCACTTGCT 335
DB 415 TTAATCAAGTAGAGTAGAGTAGAGTAAGAAATTTAAGAAAGATTCATCACTTGCT 474
QY 336 TGAACGCTGTCAGTTTCATCAGGTGATTATACCTTTGACCGGAATGTTATCCAGGC 395
DB 475 TGAACGCTGTCAGTTTCATCAGGTGATTATACCTTTGACCGGAATGTTATCCAGGC 534
QY 396 TGTAAATGAATGACAGAGATGCTGACCAATCACTTCAGCGGCCACTCAGCCCAAGT 455
DB 535 TGTAAATGAATGACAGAGATGCTGACCAATCACTTCAGCGGCCACTCAGCCCAAGT 594
QY 456 CACATGAGAGGGGTTAAATTAATGCTTTGATCATTTTTCAGATTTTGGCTGCCCT 515
DB 595 CACATGAGAGGGGTTAAATTAATGCTTTGATCATTTTTCAGATTTTGGCTGCCCT 654
QY 516 TGTATATCTTTTGGGAATTTAAACCCCACTTACAAAACATATGATGATGATCAACA 575
DB 655 TGTATATCTTTTGGGAATTTAAACCCCACTTACAAAACATATGATGATGATCAACA 714
QY 576 AAATGTTGATGAAGAAGAAATATGA 601
DB 715 AAATGTTGATGAAGAAGAAATATGA 740

```

RESULT 9
ADRI4216
ID ADRI4216 standard; DNA; 2003 BP.
XX AC ADRI4216;
XX

DT 21-OCT-2004 (first entry)

DE Human NF-kappaB pathway-associated gene SegID217.

XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
XX antiarthritic; antirheumatic; gastrointestinal-Gen; antiaesthetic;
KW antiatherosclerotic; immunomodulator; cerebroprotective; vasotrophic;
KW immunosuppressive; vulnary; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW hepatic disorder; Hodgkin's lymphoma; hematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW viral replication; host cell survival; evasion of immune response;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; gene; ds; human.

XX Homo sapiens.

XX MO2004065577-A2.

XX 05-AUG-2004.

XX 13-JUN-2004; 2004MO-US000798.

XX 14-JAN-2003; 2003US-0440068P.

PR 12-MAY-2003; 2003US-0469757P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Nadler SG, Neubauer MG, Feder JN, Carman J;

DR WPI; 2004-562168/54.

XX P-PSDB; ADRI4217.

PT New isolated polynucleotides and polypeptides associated with NF-kappaB

PT pathway, useful for diagnosing, treating, or preventing disorders or

XX diseases associated with NF-kappaB pathway.

PS Claim 1; SEQ ID NO 217; 237bp; English.

XX This invention relates to the novel association of protein sequences (and
CC the genes which encode them) to the NF-kappaB pathway. The invention may
CC be useful for the production of compounds with an antiinflammatory,
CC cytostatic, hepatotropic, virucide, antiaesthetic, antirheumatic,
CC gastrointestinal-gen, antiaesthetic, antiatherosclerotic,
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC vulnary activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappaB pathway. The condition is an
CC immune disorder, an inflammatory disorder, an inflammatory disorder
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, hematopoietic tumours, hyper-IgM
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction,
CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human gene which is
CC subject to the novel association with the NF-kappaB pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.

XX Sequence 2003 BP; 667 A; 324 C; 406 G; 606 T; 0 U; 0 Other;
SQ
Query Match 93.9%; Score 564.4; DB 13; Length 2003;
Best Local Similarity 99.8%; Pred. No. 1.6e-156;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 36 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 95
DB 205 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 264
QY 96 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACAGACAAGTAGTGGTGC 155
DB 265 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACAGACAAGTAGTGGTGC 324
QY 156 TGGATGAGCTCTACAGAGTACAGAGGAGTACACCCAAAACAAGAGAGGACAGAGA 215
DB 325 TGGATGAGCTCTACAGAGTACAGAGGAGTACACCCAAAACAAGAGAGGACAGAGA 384
QY 216 TCATCAAGAACCTCATCAAGACAGCTCATCAAGCTGGCCATTTCTTATAGAAATATCA 275
DB 385 TCATCAAGAACCTCATCAAGACAGCTCATCAAGCTGGCCATTTCTTATAGAAATATCA 444
QY 276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAGAGAAAGTTCATCAGCTTGCTA 335
DB 445 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAGAGAAAGTTCATCAGCTTGCTA 504
QY 336 TGAACGCTGTCAGTTTCATCAGGTGATTAACCTTGAACCGAATGCTTAATCCAGGC 395
DB 505 TGAACGCTGTCAGTTTCATCAGGTGATTAACCTTGAACCGAATGCTTAATCCAGGC 564
QY 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCATCGCAAGT 455
DB 565 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCATCGCAAGT 624
QY 456 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGGAATTTTGGCTGCTT 515
DB 625 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGGAATTTTGGCTGCTT 684
QY 516 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGTGATGATCAACA 575
DB 685 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGTGATGATCAACA 744
QY 576 AAATGTTGATGAAGAGAAATATGA 601
DB 745 AAATGTTGATGAAGAGAAATATGA 770
RESULT 10
ACN44179
ID ACN44179 standard; cDNA; 2034 BP.
XX ACN44179;
AC ACN44179;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human mRNA sequence HCT1950131.
XX
KM Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX
OS Homo sapiens.
XX
XX WO2003073826-A2.
XX
XX PD 12-SEP-2003.
XX
XX PF 28-FEB-2003; 2003WO-US006235.
XX
XX PR 01-MAR-2002; 2002US-00087192.
XX
XX PA (SAGR-) SAGRES DISCOVERY.
XX
XX PI Morris DW;

XX
DR WPI; 2003-328604/31.
XX
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 497; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA gene
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 2034 BP; 639 A; 352 C; 412 G; 631 T; 0 U; 0 Other;
Query Match 93.9%; Score 564.4; DB 11; Length 2034;
Best Local Similarity 99.8%; Pred. No. 1.7e-156;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 36 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 95
DB 249 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 308
QY 96 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACAGACAAGTAGTGGTGC 155
DB 309 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACAGACAAGTAGTGGTGC 368
QY 156 TGGATGAGCTCTACAGAGTACAGAGGAGTACACCCAAAACAAGAGAGGACAGAGA 215
DB 369 TGGATGAGCTCTACAGAGTACAGAGGAGTACACCCAAAACAAGAGAGGACAGAGA 428
QY 216 TCATCAAGAACCTCATCAAGACAGCTCATCAAGCTGGCCATTTCTTATAGAAATATCA 275
DB 429 TCATCAAGAACCTCATCAAGACAGCTCATCAAGCTGGCCATTTCTTATAGAAATATCA 488
QY 276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAGAGAAAGTTCATCAGCTTGCTA 335
DB 489 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAGAGAAAGTTCATCAGCTTGCTA 548
QY 336 TGAACGCTGTCAGTTTCATCAGGTGATTAACCTTGAACCGAATGCTTAATCCAGGC 395
DB 549 TGAACGCTGTCAGTTTCATCAGGTGATTAACCTTGAACCGAATGCTTAATCCAGGC 608
QY 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCATCGCAAGT 455
DB 609 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCATCGCAAGT 668
QY 456 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGGAATTTTGGCTGCTT 515
DB 669 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGGAATTTTGGCTGCTT 728
QY 516 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGTGATGATCAACA 575
DB 729 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGTGATGATCAACA 788
QY 576 AAATGTTGATGAAGAGAAATATGA 601
DB 789 AAATGTTGATGAAGAGAAATATGA 814
RESULT 11
ADRI4038
ID ADRI4038 standard; DNA; 2081 BP.

XX ADR14038;
 XX
 XX 21-OCT-2004 (first entry)
 XX
 DE Human NF-kappaB pathway-associated gene Segid39.
 XX
 XX NF-kappaB pathway; antiinflammatory; cytosolic; hepatotropic; virucide;
 KW antirheumatic; antirheumatic; gastrointestinal-Gen; antiaesthetic;
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnery; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; gene; ds; human.
 XX
 OS Homo sapiens.
 XX
 PN MO2004065577-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004MO-US000798.
 XX
 PR 14-JAN-2003; 2003JUS-0440068P.
 XX
 PR 12-MAY-2003; 2003JUS-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Nadler SG, Neubauer MG, Feder JN, Carman J;
 XX
 DR WPI, 2004-562168/54.
 XX
 DR P-PSDB; ADR14039.
 XX
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 PS Claim 1; SEQ ID NO 39; 237pp; English.
 XX
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytosolic, hepatotropic, virucide, antirheumatic, antipneumatic,
 CC gastrointestinal-Gen, antiaesthetic, antiarteriosclerotic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnery activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, cancer, aberrant apoptosis,
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
 CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction,
 CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human gene which is

CC subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.

XX SQ Sequence 2081 BP; 668 A; 359 C; 422 G; 632 T; 0 U; 0 Other;

Query Match 93.9%; Score 564.4; DB 13; Length 2081;
 Best Local Similarity 99.8%; Pred. No. 1.7e-156;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCGTTCAGGCACAAAAAGAAATCTTGG 95
 DB 268 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCGTTCAGGCACAAAAAGAAATCTTGG 327
 QY 96 GTTAAATGCTGTCCAAATCCATGCGCACACCTTAATGAGACACCAATGATGAGTGC 155
 DB 328 GTTAAATGCTGTCCAAATCCATGCGCACACCTTAATGAGACACCAATGATGAGTGC 387
 QY 156 TGGATGAGCTTACAGAGTACAGGAGTACACCCAAACAAAGAGAGCAGAGAGA 215
 DB 388 TGGATGAGCTTACAGAGTACAGGAGTACACCCAAACAAAGAGAGCAGAGAGA 447
 QY 216 TCATCAGAACTCATCAAGACAGTCATCAGCTGCGCATCTTTATAGAAATATCACT 275
 DB 448 TCATCAGAACTCATCAAGACAGTCATCAGCTGCGCATCTTTATAGAAATATCACT 507
 QY 276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCGCTTCTA 335
 DB 508 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCGCTTCTA 567
 QY 336 TGACGTGCTCAGTTTCCATCAGTGTGATTTATACCTTTGACCGGATGTTATCCAGGC 395
 DB 568 TGACGTGCTCAGTTTCCATCAGTGTGATTTATACCTTTGACCGGATGTTATCCAGGC 627
 QY 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCCACCTCACTGCCAGT 455
 DB 628 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCCACCTCACTGCCAGT 687
 QY 456 CACATGAGACGGGTAAATATGCTTGAATCATTTTTCAGATGTGAAATTTTGGCGCT 515
 DB 688 CACATGAGACGGGTAAATATGCTTGAATCATTTTTCAGATGTGAAATTTTGGCGCT 747
 QY 516 TGTAAATCCTTTTGGGAATTTAAACCCCACTTCAAAAACATATGATGATCAACA 575
 DB 748 TGTAAATCCTTTTGGGAATTTAAACCCCACTTCAAAAACATATGATGATCAACA 807
 QY 576 AATGTTGATGAAGAGAAATATGA 601
 DB 808 AATGTTGATGAAGAGAAATATGA 833
 RESULT 12
 ACN44178
 ID ACN44178 standard; DNA; 58723 BP.
 XX
 ACN44178;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 XX Human genomic sequence hCG36837.
 DE
 XX
 KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
 OS
 XX Homo sapiens.
 XX
 PN WO2003073826-A2.
 XX
 PD 12-SEP-2003.
 XX
 PF 28-FEB-2003; 2003WO-US006235.
 XX
 PR 01-MAR-2002; 2002US-00087192.
 XX

PA (SAGR-) SAGRES DISCOVERY.
 XX
 PI Morris DW;
 XX WPI; 2003-328604/31.
 XX
 DR
 XX
 PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.
 XX
 XX Claim 1; SEQ ID NO 496; Opp; English.
 XX
 XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcino Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bioclip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcino Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published
 CC
 XX
 XX Sequence 58723 BP; 16691 A; 11477 C; 12539 G; 17816 T; 0 U; 0 Other;
 SQ
 Query Match 93.9%; Score 564.4; DB 11; Length 58723;
 Best Local Similarity 99.8%; Pred. No. 8,9e-156;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 95
 DB 46938 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 46997
 QY 96 GTAAATGCTGTCCTCAATTCATGCCCACTTAATAGACACAAAGTAGAGGTGC 155
 DB 46998 GTAAATGCTGTCCTCAATTCATGCCCACTTAATAGACACAAAGTAGAGGTGC 47057
 QY 156 TGGATGAGCTCTACAGAGTACAGGAGGATACACCCAAAACAGAGAGGACAGAGA 215
 DB 47058 TGGATGAGCTCTACAGAGTACAGGAGGATACACCCAAAACAGAGAGGACAGAGA 47117
 QY 216 TCATCAAGAACTCATCAAGACAGTCAATGAGTGGCCATCTTTATAGGAATATCAGT 275
 DB 47118 TCATCAAGAACTCATCAAGACAGTCAATGAGTGGCCATCTTTATAGGAATATCAGT 47177
 QY 276 TTATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGTCTA 335
 DB 47178 TTATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGTCTA 47237
 QY 336 TGACCGTGTGCTGATTCATCAGGTGATATACCTTTGACGGGAATGTTATCCAGGC 395
 DB 47238 TGACCGTGTGCTGATTCATCAGGTGATATACCTTTGACGGGAATGTTATCCAGGC 47297
 QY 396 TGTAAATGATGACAGAGATGCTGACCAATTCATTCAGCGCCACCTCATGCCAAGT 455
 DB 47298 TGTAAATGATGACAGAGATGCTGACCAATTCATTCAGCGCCACCTCATGCCAAGT 47357
 QY 456 CACATGAGCGGGTTAATATGCTTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
 DB 47358 CACATGAGCGGGTTAATATGCTTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 47417
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATATACACA 575
 DB 47418 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATATACACA 47477
 QY 576 AAATGTTGATGAGAGACATATGA 601
 DB 47478 AAATGTTGATGAGAGACATATGA 47503

RESULT 13
 ID AAH07403
 ID AAH07403 standard; cDNA; 816 BP.
 XX
 AC AAH07403;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:4238.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 1; SEQ ID NO 4238; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SO Sequence 816 BP; 262 A; 155 C; 180 G; 212 T; 0 U; 7 Other;
 Query Match 91.1%; Score 547.6; DB 4; Length 816;
 Best Local Similarity 98.9%; Pred. No. 1e-151;
 Matches 562; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
 QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 95
 DB 102 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 161

QY 96 GTAATAATGGTGTCCAAATCCATGCGCCACCCTTAATAGACGACAGATGAGGTGC 155
 DB 162 GTTAAATGGTGTCCAAATCCATGCGCCACCCTTAATAGACGACAGATGAGGTGC 221
 QY 156 TGGATGAGCTTACAGAGTACGACGAGAGTACCCAAAACAAGAGGAGAGAGA 215
 DB 222 TGGATGAGCTTACAGAGTACGACGAGAGTACCCAAAACAAGAGGAGAGAGA 281
 QY 216 TCATCAAGAACCTTCATCAAGACAGTCAAGAGTGCACCTTTTATAGAAATACAT 275
 DB 282 TCATCAAGAACCTTCATCAAGACAGTCAAGAGTGCACCTTTTATAGAAATACAT 341
 QY 276 TTAATCAAGATGAGTACGATGATGAGAAATTTAAGAAAGATTCACCTTGCTA 335
 DB 342 TTAATCAAGATGAGTACGATGATGAGAAATTTAAGAAAGATTCACCTTGCTA 401
 QY 336 TGAACGTGTCACTTTCCATCAGTGTGATTAACCTTTGACCGAAATGTGTATCCAGC 395
 DB 402 TGAACGTGTCACTTTCCATCAGTGTGATTAACCTTTGACCGAAATGTGTATCCAGC 461
 QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATGAGCGCCACCTTCACCTCAAGT 455
 DB 462 TGTAAATGAATGACAGAGATGCTGCACCAATTCATGAGCGCCACCTTCACCTCAAGT 521
 QY 456 CACATGACGAGGTTAATATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCCT 515
 DB 522 CACATGACGAGGTTAATATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCCT 581
 QY 516 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAAC 575
 DB 582 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAAC 641
 QY 576 AAT--GTGATGAGAGAGAACTATGA 601
 DB 642 AATGTGTTGATGAAGAGAACTATGA 669

RESULT 14

ACH89719
 ID ACH89719 standard; DNA; 544 BP.

XX ACH89719;

DT 29-JUL-2004 (first entry)

DE Human genome derived single exon probe #22914.

KW Human; probe; ss; gene expression; single exon probe; microarray;

XX alternative splicing event; genomic alteration.

OS Homo sapiens.

PN US2003194704-A1.

PD 16-OCT-2003.

PF 03-APR-2002; 2002US-00029386.

PR 03-APR-2002; 2002US-00029386.

PA (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZ/) HANZEL D K.

PI Penn SG, Rank DR, Hanzel DK;

DR WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.

PS Claim 1; SEQ ID NO 22914; 80bp; English.
 XX The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX

SO Sequence 544 BP; 175 A; 109 C; 117 G; 143 T; 0 U; 0 Other;

Query Match 90.2%; Score 542.4; DB 12; Length 544;
 Best Local Similarity 99.8%; Pred. No. 2.9e-150;

Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTTCCAAAACCTGCGCGTTCAGGCACAAAAAGAAATCTTGG 95
 DB 1 TGGCCACAGATGCTTTAATTTCCAAAACCTGCGCGTTCAGGCACAAAAAGAAATCTTGG 60
 QY 96 GTAATAATGGTGTCCAAATCCATGCGCACCACTTAATAGACGACAAATAGTAGGTGC 155
 DB 61 GTAATAATGGTGTCCAAATCCATGCGCACCACTTAATAGACGACAAATAGTAGGTGC 120
 QY 156 TGGATGAGCTTACAGAGTACAGGAGTACACCCAAAACAAGAGGAGAGAGA 215
 DB 121 TGGATGAGCTTACAGAGTACAGGAGTACACCCAAAACAAGAGGAGAGAGA 180
 QY 216 TCATCAAGAACTTCATCAAGAGATCATACACTGCCCTTTTATAGAAATATCACT 275
 DB 181 TCATCAAGAACTTCATCAAGAGATCATACACTGCCCTTTTATAGAAATATCACT 240
 QY 276 TTAATCAAGATGAGTACGATGATGAGAAATTTAAGAAAGATTCACCTTGCTGA 335
 DB 241 TTAATCAAGATGAGTACGATGATGAGAAATTTAAGAAAGATTCACCTTGCTGA 300
 QY 336 TGAACGTGTCACTTTCCATCAGTGTGATTAACCTTTGACCGGAATGTGTATCCAGC 395
 DB 301 TGAACGTGTCACTTTCCATCAGTGTGATTAACCTTTGACCGGAATGTGTATCCAGC 360
 QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTTAGAGGCCACCTTCACCTCAAGT 455
 DB 361 TGTAAATGAATGACAGAGATGCTGCACCAATCATTTAGAGGCCACCTTCACCTCAAGT 420
 QY 456 CACATGACGAGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCCT 515

Db 421 CACATGACGGGTTAAATATCTGTTGATCATTTTTCAGATGTGGAATTTTGGCTCCT 480
 QY 516 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGTATCAACA 575
 Db 481 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGTATCAACA 540
 QY 576 AAAT 579
 Db 541 AAAT 544

RESULT 15

ACN44176
 ID ACN44176 standard; DNA; 62231 BP.

ACN44176;

DT 18-NOV-2004 (first entry)

XX Mouse genomic sequence MCG15520.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX Mus musculus.

PN WO2003073826-A2.

PD 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006235.

PR 01-MAR-2002; 2002US-00087192.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW;

DR WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 493; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bioclip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is an equivalent to basic patent
 CC US0202182586A1, for which no sequence data was published

XX Sequence 62231 BP; 15482 A; 12591 C; 12729 G; 16514 T; 0 U; 4915 Other;

Query Match 74.2%; Score 446; DB 11; Length 62231;

Best Local Similarity 86.1%; Pred. No. 1.2e-120;

Matches 494; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 28 TCTCCGATGCGCACAGATCTTTTAATTCACAAAACCTGCGCGTTCAGGCACAAAAAGAA 87

Db 50644 TTCTGACAGTGGCTACAGATGTTTCATTTCAAAAACCTGCGCGTTCAGGCACAAAAAGAA 50703

QY 88 GATCTTGGGTAATAGTGTCCAAATCCATGCGCACACCTTAATAGACAGACAGTAG 147

Db 50704 GATCTTGGGCAAAATGTAATCCAAATCCATGCGCACACCTTAATAGACAGACAGTAG 50763

QY 148 TGAAGTCTGATGAGCTCTACAGATGACACAGGAGTACCCCAAAAAGAAAGAGGC 207
 Db 50764 CGAGGTGCTGATGAGCTGTATACAGGAGTACCAAGAGGATACCCAGAAAGAAAGAGGC 50823
 QY 208 AGAAGATGATGAAAGAACCTCATCAAGACAGATCAACAGCTGGCCATTCTTATAGAA 267
 Db 50824 GGAAGAGGTATCAAGAACCTCATCAAGACAGATCAACAGCTGGCCATTCTTCAACAGAA 50883
 QY 268 TAATCAGTTTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTATCA 327
 Db 50884 CAATCAGTTTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTATCA 50943
 QY 328 GCTTGCTATGACCGGTGTCAGTTTCCATCAAGTGTGATTAATCTTGAACCGGAATGTGTT 387
 Db 50944 GCTTGCTATGACCGGTGTCAGTTTCCATCAAGTGTGATTAATCTTGAACCGGAATGTGTT 51003
 QY 388 ATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATGATTCAGGCGCACTTCAC 447
 Db 51004 GTCCAGGCTGCTGAACGAGTGCAGAGCTCTTACACGAGATATTCAGGCGCACTTCAC 51063
 QY 448 TGCCAGTGCATGACGAGGTTAATATGCTTTGATCATTTTTCAGATTGTAATTTT 507
 Db 51064 CGCCAACTGTCACGAGACGGTTAATATGCTTTGACCATTTTTCAGATTGTAATTTT 51123
 QY 508 GCGTGCCTGTATATATCTTTTGGAAATTTTAAACCCCACTTACAAAACCTATGTGATG 567
 Db 51124 GCGTGCCTGTATATATCTTTTGGAAATTTTAAACCTTCACTTACAAAACCTTGGCAGCG 51183
 QY 568 TATCAACAAAATGTTGATGAGAAAGAAATATCA 601
 Db 51184 CATCAACAAAATGTTGATGAGAAAGAAATATCA 51217

RESULT 16

ACN44177
 ID ACN44177 standard; cDNA; 2087 BP.

ACN44177;

DT 18-NOV-2004 (first entry)

XX Mouse mRNA sequence MCT17722.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.

XX Mus musculus.

PN WO2003073826-A2.

PD 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006235.

PR 01-MAR-2002; 2002US-00087192.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW;

DR WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 494; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

Db 405 TGGGGAATTTAGAAAGTTTCATCAGCTGTGTATGACCGTGGNAGTTCCAAACAG 464
 QY 360 TGGATTATACCTTTGACCGGATGTGTTATTCAGCGCTGTTAAATGAATGCAG-AGAGATG 418
 Db 465 GGGGTTATACCTTTGACCGGATGTGTTATTCAGCGCTGTTAAATGAATGCANNAAGAAATG 524
 QY 419 CTGACCAATTCATCA 435
 Db 525 CTGACCAAAACANTTA 541

RESULT 18

ADQ63108
 ID ADQ63108 standard; cDNA; 1986 BP.

AC ADQ63108;

DT 07-OCT-2004 (first entry)

DE Novel human cDNA sequence #269.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
 KM cytotactic; gene therapy; diagnostic marker; morbid state; osteoporosis;
 KM neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
 KM cancer.

XX Homo sapiens.

XX EP1440981-A2.

XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

XX 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.

XX P-PSDB; ADQ65236.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 269; 2449BP; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.

XX Sequence 1986 BP; 536 A; 458 C; 497 G; 495 T; 0 U; 0 Other;

XX Query Match 42.3%; Score 254.4; DB 12; Length 1986;

XX Best Local Similarity 65.9%; Pred. No. 1.6e-64;

XX Matches 369; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 42 CAGATGCTTTTATTCAGAAACCTGGCCCTTACGACACAAAGAGATTTGGTAAAA 101

Db 150 CTGATGCTTTTATTCAGAAAGCTTTCGCTTCAAGCCGAGAGAGATTCGACAAAA 209

QY 102 TGGTGTCCAATCATCGCACACCTTATATAGACGACAAAGATGAGTGCTGTGATG 161

Db 210 TAGCAGCAAAAGTGTGGCCACATGTGATGTATGACACACAGACGAGATCTTGATG 269
 QY 162 AGCTTACAGATGACGACGAGATGACCCAAAACAAAGAGAGGACGAGATCATCA 221
 Db 270 AGCTTACAAAGTACACCAAGACACACACAAAGAGAGAGCCACAGATCATCA 329
 QY 222 AGAACCTCATCAGACAGTATCAGTGTGCGCCATCTTTTATAGGAATATCAGTTTATC 281
 Db 330 AAGACTTAATCAAGGTGGCGATCAAAATCGGATCTCTACCGAATACACAGTTTATGCC 389
 QY 282 AAGATGAGCTAGCATTTATGAGAAATTTAAGAAAGATTCATCAGCTTGTATGACCG 341
 Db 390 AAGAGAGAGCTGTTATTTGTGAGAAAGTCCGGAAGAGCTGAAACCGCATGACCA 449
 QY 342 TGGTCAAGTTTCATCAGGTGATTTATCCTTTACCGGAAATGTGTTATCAGGCTGTAA 401
 Db 450 TTGTCAAGCTTCTATGAGGTGATATACCTTTCATGAGAAAGTGTCTCTCCATCTCTGC 509
 QY 402 ATGAATGACAGAGATGCTGACCAAAATCATTTAGCGCCACCTCATCTGCCAAGTCATG 461
 Db 510 ATAGTGCAGAGACCTGTGTGATGAATCTGTGACGCGCACCTGACGCCACGACCCACG 569
 QY 462 GACGGGTTAATATGCTTTGATCATTTTTCAGATTTGTAATTTTGGCTGCTGTATA 521
 Db 570 GGGGCTATCAACACGCTTTTACCACTTTGCGATGTGAGATTCCTTCCACCTCTATA 629
 QY 522 ATCTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAAAAATGT 581
 Db 630 GTCTGATGAGACTGTGAGCCCAACCTCAAGAGATTTGTGAAGATCATATATAGTTC 689
 QY 582 TGGATGAAGAAACATATGA 601
 Db 690 TAGATGAGAAAGTCTTTTAA 709

RESULT 19

ADQ30769
 ID ADQ30769 standard; cDNA; 969 BP.

XX ADC30769;

XX 18-DEC-2003 (first entry)

XX Human novel cDNA sequence, SEQ ID NO:851.

XX Human; diagnostic; drug screening; forensics; gene mapping;
 KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KM ulcers; osteoporosis; autoimmune disease; cancer;
 KM molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KM neuroprotective; antiplatelet; anticoagulant; thrombolytic; vulnerary;
 KM anticancer; osteopathic; immunosuppressive; antiinflammatory; cytotactic;
 KM gene therapy; chromosome 15q21.2; gene; ss.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;
 PI Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

XX P-PSDB; ADQ31740.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.

PS Claim 1; SEQ ID NO 851; 1185bp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX invention also relates to nucleic acid sequences over 98% identical with
XX the novel human cDNAs. The invention additionally encompasses expression
XX vectors and host cells comprising a nucleic acid of the invention; the
XX recombinant production of a polypeptide of the invention; an antibody
XX against a polypeptide of the invention; a method of detecting
XX polynucleotides or polypeptides of the invention; and methods of
XX identifying a compound which binds to a polypeptide of the invention. The
XX invention further discloses methods of preventing, treating or
XX ameliorating a medical condition; kits comprising polynucleotide probes
XX and/or monoclonal antibodies for carrying out the methods of the
XX invention; methods for the identification of compounds that modulate the
XX expression or activity of the polynucleotide and/or polypeptide; and 767
XX coding sequences corresponding to the cDNA sequences of the invention
XX (ADC31861-ADC32627) and the polypeptides encoded by the cDNAs (ADC32628
XX -ADC33394). The nucleic acids and polypeptides of the invention are
XX useful in diagnostics, drug screening, forensics, gene mapping, in the
XX identification of mutations responsible for genetic disorders or other
XX traits, for assessing biodiversity, and in producing many other types of
XX data and products dependent on DNA and amino acid sequences. They are
XX also used for treating diseases such as Parkinson's disease, Alzheimer's
XX disease and other neurodegenerative diseases, anaemia, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer. The nucleic acids may also be used as hybridisation probes or
XX primers, and in the recombinant production of a protein. The polypeptides
XX are also useful in generating antibodies, as molecular weight markers,
XX and as food supplements. The present sequence represents a specifically
XX claimed human cDNA sequence of the invention. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 969 BP; 264 A; 261 C; 257 G; 187 T; 0 U; 0 Other;

Query Match 41.6%; Score 249.8; DB 10; Length 969;
Best Local Similarity 66.4%; Pred. No. 2,6e-63;
Matches 359; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 42 CAGATGCTTTTAATTCACAAAACCTGGCCGTTCAGGCACAAAAGAGATCTGGGTAATA 101
DB 320 CTGATGCTTTTAATTCACAAAAGCTTGGCGCTTCAGCCACGAAAGATCTGACGAAAA 379
QY 102 TGGTGTCCAAATCCATCGCACCACTTAATGAGACACAACTAGTGGGTGGATG 161
DB 380 TAGCCAGCAAAATGTGGCCAAATGTGATGATATACACGACACGAGATCTTGGATG 439
QY 162 AGCTTACAGATGACGACGAGGAGTACACCCAAAACAAGAGAGGAGAGAGATCATCA 221
DB 440 AGCTTACAAAAGTCAACCAAAAGCACACACAAACAAGAGAGGCCACAAATCATGA 499
QY 222 AGAATCTCATCAAGACAGTCATAGCTGGCCATTTTATAGAAATTAATCACTTTAATC 281
DB 500 AAGACTTAATCAAGGTGGGAGTCAAAATGGGATCCTCTACCGAAACAACAGTTTATGACC 559
QY 282 AAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGAAAGTCACTACGTTGCTATACCG 341
DB 560 AAGAGAGAGCTGTTTATTTGAGAAAGTCCGGAAAGAGCTGAACCGACCGCATACCA 619
QY 342 TGGTCAGTTTTCATCAGTGGATTAATACCTTTGACCGGATGATGTTATCCAGGCTGTAA 401
DB 620 TTGTGAGCTTTTATGAGGTGAATACCTTGGATGGAACGTCCTCCAAATCTCTGCG 679
QY 402 ATGAATGACAGAGATGCTGACCAAAATCACTTACGGCCCACTCACTGCCAATGCAATG 461
DB 680 ATGAGTGAAGGACCTGTGTCATGAATCTGGTGCAGGCGGACCTGAGCGCCAGAGACCACG 739

QY 462 GACGGTTAATATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATA 521
DB 740 GGGCATCAACCAAGCTCTTTTAACACCTTGGCGATGTGAGATTCCTCTCCACCTCTATA 799

QY 522 ATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTAGATGATCAACAAAATGT 581
DB 800 GTCTGATGAGAGACTGTAGGCCCACTCAAGAGATTTGTGAAGAAATCAATATAGTTGC 859

QY 582 T 582
DB 860 T 860

RESULT 20

ACH92216
ID ACH92216 standard; DNA; 527 BP.

ACH92216;

29-JUL-2004 (first entry)

Human genome derived single exon probe #25411.

Human; probe; ss; gene expression; single exon probe; microarray;
alternative splicing event; genomic alteration.

Homo sapiens.

US2003194704-A1.

16-OCT-2003.

03-APR-2002; 2002US-00029386.

03-APR-2002; 2002US-00029386.

(PENN/) PENN S G.

(RANK/) RANK D R.

(HANZ/) HANZEL D K.

Penn SG, Rank DR, Hanzel DK;

WPI; 2004-119264/12.

New human genome-derived single exon nucleic acid probes useful for human
gene expression analysis, for identifying or characterizing alternative
splicing events, for assessing genomic alterations or as tools for
surveying tissues.

Claim 1; SEQ ID NO 25411; 80bp; English.

The invention relates to a nucleic acid probe for measuring human gene
expression, comprising any of the 27,400 fully defined nucleotide
sequences in the specification, or their complements or fragments, and
encoding at least 8 amino acids of any of the 688 amino acid sequences
fully defined in the specification. The probe is a single exon probe that
hybridises under high stringency conditions to a nucleic acid molecule
expressed in human cells or tissues. Also included are a spatially-
addressable set of single exon nucleic acid probes for measuring human
gene expression (comprising a plurality of single exon nucleic acid
probes cited above, where each of the plurality of probes is separately
and addressably isolatable or amplifiable from the plurality), a single
exon microarray for measuring human gene expression, a method of
measuring human gene expression, a vector comprising the single exon
probe cited above, an ORF-encoded peptide comprising at least 8
contiguous amino acids of any of the above-mentioned amino acid
sequences (optionally with conservative amino acid substitutions), an
isolated antibody that binds specifically to a peptide cited above,
methods of selling and/or licensing single exon probes or microarrays to
a customer desiring to measure gene expression, a method of providing
human gene expression data by subscription, and a computer-readable
storage medium which contains a database having a plurality of records

CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPRO at
CC seqdata.uspro.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 527 BP; 163 A; 129 C; 128 G; 107 T; 0 U; 0 Other;
Query Match 39.5%; Score 237.4; DB 12; Length 527;
Best Local Similarity 65.7%; Pred. No. 9.1e-60;
Matches 346; Conservative 0; Mismatches 181; Indels 0; Gaps 0;
QY 75 AGGCACAAAAGAGATCTTGCTAAATGCTCAATCCATCCGCCACCTTAATAG 134
Db 1 AAGCCAGAGAGAGATCTGAGCAAAATAGCCAGCAAACTGTGGCCACATGTTGATTG 60
QY 135 ACGACACAGTATGAGTGTGATGAGCTTTCACAGAGTGACCGAGGAGTACACCCAA 194
Db 61 ATGACACACAGCAGCATCTTTGATGAGCTTCAAAAGTCCACCAAGACACACACACA 120
QY 195 ACAAGAGAGGAGGAGAGATCATCATCAAGAACCTCATCAGACAGCTCATAGCTGCCA 254
Db 121 ACAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 255 TTTCTTATAGGAATTAATCACTTTAATCAAGATGAGTATGATGAGGAGGAGGAGGAG 314
Db 181 TCTCTACCGGAGACACATTTAGCCAGAGAGAGCTGTTATTTGAGAGAGTCCGGA 240
QY 315 AGAAGATTATCACTGCTGTATAGCCGTGTCACTTTCATCAGTGGATTAATACCTTGG 374
Db 241 AGAAGCTGAACCAAGCCGCGCATGACCTTGTGAGTGAATTAACCTTGG 300
QY 375 ACCGGAATGTTATCCAGGCTGTAAATGAATGAGAGAGATGCTGACCAATATCTTC 434
Db 301 ATAGGAACGCTGCTCTCAATCTCTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 435 AGCGCACCTCACTGCAATGATCAGATGAGCGGTTAATATGCTTTGATCATTTTTCAG 494
Db 361 AGCGGACCTGAGCGCCAGAGACCCAGCGGCGCATCAACGCTTTTAACCATTTGGCG 420
QY 495 ATTGTGAATTTTGGCTGCTGCTGTATATCTTTTGGGAATTTTAAACCCACTTCAAA 554
Db 421 ATGTGAGTTCCTCTCCACCTCTATATGCTGTGATGAGGAGCTGAGGCCCACTCAAGA 480
QY 555 AACTATGATGTGTATCAAAAATGTTGATGAAGAAACATATGA 601
Db 481 GGATTTGTGAAGAAATCAATAGTTGCTAGATGAGAAAGTCTTTAA 527
RESULT 21
ABS78741
ID ABS78741 standard; DNA; 645 BP.
XX
AC ABS78741;
XX
DT 16-DEC-2002 (first entry)
XX
DE DNA encoding human NOVX17a protein.
XX
KW Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
KW Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
KW tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
KW Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;

KW obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
KW infertility; inflammatory bowel disease; atherosclerosis; hypertension;
KW scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
KW achia; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
KW bacterial infection; parasitic infection; graft-versus-host disease;
KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
KW angiogenesis; gene; de.
XX
OS Homo sapiens.
XX
PN WO200272770-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US007283.
XX
PR 08-MAR-2001; 2001US-0274281P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0279366P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-MAY-2001; 2001US-0288148P.
PR 31-MAY-2001; 2001US-0294821P.
PR 31-OCT-2001; 2001US-0335302P.
PR 04-DEC-2001; 2001US-0338375P.
PR 07-MAR-2002; 2002US-00094466.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Vernet CA, Tchermey VT, Malyankar UM, Gerlach VL;
PI Li L, Zernusen BD, Patutajan M, Gusev VY, Kekuda R, Pena CBA;
PI Zhong M, Gangoli EA, Taupier RJ;
XX
DR WPI; 2002-713508/77.
XX
PT P-PSDB; ABG97497.
XX
PT New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
PT sclerosis, atherosclerosis, cancer, infections, osteoporosis or
PT Parkinson's disease.
XX
PS Claim 22; Page 139; 266pp; English.
XX
CC The present invention relates to a new polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease,
CC preferably a NOVX-associated disorder. The NOVX nucleic acids,
CC polypeptides and antibodies are useful for treating, preventing or
CC diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
CC syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, inflammatory bowel disease, atherosclerosis, hypertension,
CC scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
CC asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
CC infections, or graft-versus-host disease. The nucleic acids and
CC polypeptides may also be used as targets for the identification of small
CC molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and

CC angiogenesis, in gene therapy, in generation of antibodies that bind
 CC immunospecifically to NOXV substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids are further used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine, and
 CC pharmacogenomics. The present nucleic acid sequence encodes a human NOXV
 CC protein of the invention

XX Sequence 645 BP; 97 A; 228 C; 219 G; 101 T; 0 U; 0 Other;

Query Match 34.6%; Score 207.8; DB 6; Length 645;
 Best Local Similarity 59.9%; Pred. No. 6.1e-51;
 Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGCGCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTGAGGACAA 82
 DB 8 CTGTGTCCCGAGGCGCATGACACCTTCAGACCAAGAGCTGGCTGCGAGGGCGAG 67
 QY 83 AAGAGATCTTGGGTAAATGTTGTCGAATTCATGCGCACCACTTAAATAGACGACAA 142
 DB 68 AAGAGACTCTGTAAGATGAGTGGCTCCAAAGGACAGTGTGGCGTGTGGATGACAC 127
 QY 143 AGTAGAGAGTGTGATGATGAGCTCTACAGATGACCAAGGAGTACACCAAAACAAG 202
 DB 128 AGCAGTAGAGTGTGATGATGAGCTCTACAGATGACCAAGGAGTACACCGACCGCAAG 187
 QY 203 GAGGACAGAGATCATCAAGAACTCATCAAGACAGTCAATCAAGTGGCCATTCTTAT 262
 DB 188 GAGGCCCAAGAAATGCTCAAGAACTGCTCAAGTGGCCCTGAAGCTGGAGTGTGCTG 247
 QY 263 AGGAATTAATCATGTTAATCAAGATGAGTACAGTGAAGAAATTTAAGAAAGTT 322
 DB 248 CGTGGGACACAGCTGGGCGGTGAGAGCTGGCGCTGCGGCGCTCCGCCACCGGCG 307
 QY 323 CATCACTGTATATACCGTGTGATTTCCATCAAGTGAATTAATCTTGAACCGGAAT 382
 DB 308 CGTGTGCTGCAATGACGCGCTGACCTTCACAGTGTGACTTCACTTCGACCGGCG 367
 QY 383 GTGTATTCAGAGCTGTTAAATGATGACAGAGATCTGCAACCAATCATTCAGCGCCAC 442
 DB 368 GTGTGTGGCGCGCGGTGCTGCGAGTCCCGGACCTCTGACCAAGCCGCGGTCCCGAC 427
 QY 443 CTCACCTCCAAAGTCAATGAGCGGGTAAATATATGCTTTGATCATTTTTCAGATTGAA 502
 DB 428 CTGACCGCAAGTCCACCGGCCCATCAACCAAGTGTGGCGCACCTGACGCGAC 487
 QY 503 TTTTGGCTGCTGTATATATCCTTTTGGAAATTTAAACCCACTTACAAAACATATG 562
 DB 488 TTCTGTGGCTGCTGTATACGCGCCCGCGAGCCCTTACCGCTCCACCTGCGAGATCTGC 547
 QY 563 GATGTATCAACAAATGTTGATGAAGAGACATATGA 601
 DB 548 GAGGGCTGGGCGCGGATGCTGACGAGGAGGAGCCTCTGA 586

RESULT 22

AA139626
 ID AA139626 standard; DNA; 2108 BP.

XX AA139626;

05-SEP-2002 (first entry)

XX Human secreted protein DNA SEQ ID No 70.

DE Antiarteriosclerotic; cytoskeletal; HIV; antiallergic; antianemic;
 XX antiasthmatic; cardiast; vasotropic; neuroprotective; nootropic; SECP;
 KW anticonvulsant; antiparkinsonian; cerebroprotective; antiinflammatory;
 KW immunosuppressive; human secreted protein; cell proliferative disorder;
 KW arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;
 KW allergy; anaemia; asthma; cardiovascular disease; developmental disorder;
 KW ischaemic heart disease; congestive heart failure; neurological disorder;
 KW renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;
 KW Parkinson's disease; epilepsy; stroke; knockin humanised animal;

KW transgenic animal; gene therapy; gene; ds.

XX Homo sapiens.

OS WO200238602-A2.

XX 16-MAY-2002.

PD 08-NOV-2001; 2001WO-US047420.

XX 08-NOV-2000; 2000US-0247505P.

PR 09-NOV-2000; 2000US-0248642P.

PR 16-NOV-2000; 2000US-0249824P.

PR 21-NOV-2000; 2000US-0252824P.

PR 08-DEC-2000; 2000US-0254305P.

PR 18-DEC-2000; 2000US-0256448P.

XX (INCY-) INCYTE GENOMICS INC.

PA Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Wallia NK;
 PI Sanjwalma M, Thornton M, Elliott VS, Lu Y, Gietzen KJ, Burford N;
 PI Ding L, Hafalia AJA, Tang YT, Bandman O, Warren BA, Honchell CD;
 PI Lu DM, Thangavelu K, Lee S, Xu Y, Yang J, Lal PG, Tran B;
 PI Ison CH, Duggan BM, Saperstein SK;
 XX WPI; 2002-519296/55.
 DR P-PSDB; AAO21665.

XX Human secreted proteins and polynucleotides for diagnosing, treating or
 PT preventing disorders of cell proliferative, cardiovascular,
 PT developmental, neurological and autoimmune/inflammatory disorders.

XX Claim 5; Page 195-196; 22pp; English.

PS The invention relates to an isolated human secreted protein (SECP)
 CC polypeptide from 63 fully defined protein sequences given in the
 CC specification. The polypeptide is useful for the diagnosing/treating of a
 CC disease with decreased/overexpression of SECP. Examples of disorders
 CC associated with abnormal expression of SECP include a cell proliferative
 CC disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory
 CC disorder, AIDS, allergies, anaemia, asthma; cardiovascular disease e.g.
 CC congestive heart failure, ischaemic heart disease; developmental disorder
 CC e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.
 CC Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.
 CC The SECP polynucleotide and polypeptide are further useful for analysing
 CC the proteome of a tissue or a cell type. The polynucleotide is useful for
 CC creating knockin humanised animals (pigs) or transgenic animals (mice or
 CC rats) to model human disease, and for somatic or germ-line gene therapy,
 CC and further for generating hybridisation probes useful in mapping the
 CC naturally occurring genomic sequence. This polynucleotide sequence
 CC represents the DNA of a human secreted protein of the invention

XX Sequence 2108 BP; 408 A; 602 C; 636 G; 462 T; 0 U; 0 Other;

QY Query Match 34.6%; Score 207.8; DB 6; Length 2108;
 Best Local Similarity 59.9%; Pred. No. 1.1e-50;
 Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGCGCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTGAGGACAA 82
 DB 412 CTGTGTCCCGAGGCGCATGACACCTTCACACCAAGAGCTGTGCTGAGGCGCAG 471
 QY 83 AAGAGATCTTGGGTAAATGTTGTCGAATTCATGCGCACCACTTAAATAGACGACAA 142
 DB 472 AAGAGACTCTTAAGTAAGTGGCTGTCGAAGAGTGTGGCCGTGTGATGATGACAC 531
 QY 143 AGTAGAGAGTGTGATGATGAGCTCTACAGATGACCAAGGAGTACACCAAAACAAG 202
 DB 532 AGCAGTAGAGTGTGATGATGAGCTCTACAGATGACCAAGGAGTACACCGGACCGCAAG 591
 QY 203 GAGGACAGAGATCATCAAGAACTCATCAAGACAGTCAATCAAGTGGCCATTCTTAT 262
 DB 592 GAGGCCCAAGAAATGCTCAAGAACTGTGCAAGTGGCCCTGAAGCTGGAGTGTGCTG 651

QY 263 AGGAATAATGATTATTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGTT 322
 DB 652 CTTGGGAGCAAGCTGGCGCGGTGAGAGCTGCGCTTGGCGCGCTCCGCCACCGGCG 711
 QY 323 CATCAGCTTCTATGACCGTGTGAGTTTCCATCAGGTGATTTATCTTTGACCGGAA 382
 DB 712 CGTGGCTTGGCGATGAGCGCGCTTCCATCCAGATGAGCTTCACTTCCAGCGCGC 771
 QY 383 GTGTTATCCAGGCTGTTAAATGAATGAGAGATGCTGCAACCAATCTTACCGCGCAC 442
 DB 772 GTGCTGGCGCGCGGCTGCTGAGTGGCGGACCTGTGACACGAGCGCGTGGTCCCGAC 831
 QY 443 CTGACGCCAAGTCACATGAGACGGGTAAATATCTCTTTGATCATTTTTCAGATTGGA 502
 DB 832 CTGACCGCAAGTCCACGCGCGCATCAACCATGTTGCGCACCTTACCGCATGCGAC 891
 QY 503 TTTTGGCTGCTTGTATTAATCTTTTGGGAATTTTAAACCCCACTTAAACAACTATGT 562
 DB 892 TTTCTGGCTGCTGCTTCAAGCGCGCGCGGACCGCTTCCACCTTCCGCGAGATCTGC 951
 QY 563 GATGTTATCAACAATAATGTTGATGAGAGAAATATGA 601
 DB 952 GAGGCGCTGGCGCGGATGCTGACGAGGAGGAGCGCTTGA 990

 RESULT 23
 ID ADQ85659 standard; cDNA; 2186 BP.
 AC ADQ85659;
 DT 07-OCT-2004 (first entry)
 XX Human tumour-associated antigenic target (TAT) cDNA sequence #2473.
 DE human: tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KM cancer; cell proliferative disorder; gene; ss.
 OS Homo sapiens.
 XX MO2004060270-A2.
 PN 22-UTL-2004.
 XX 15-OCT-2003; 2003MO-US029126.
 PF 18-OCT-2002; 2002US-0418988P.
 PR (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 PI Wu TD, Zhou Y;
 XX MPI, 2004-534300/51.
 DR New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX Claim 1; SEQ ID NO 2473; 5504bp; English.
 PS
 CC The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO.1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence

CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.
 XX
 SQ Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;

 Query Match 34.0%; Score 204.2; DB 12; Length 2186;
 Best Local Similarity 59.3%; Pred. No. 1.3e-49;
 Matches 347; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

 QY 17 CGGTGGCCGACCTCCGATGCGCACAGATGCTTAAATCCAAAGCTGCGCTTAC 76
 DB 92 CGACTCTGTGTAAGATGAGCGATGACACCTTACGACCAAGAGCTGCTGCGAG 151
 QY 77 GCACAAAGAGATCTTGGGTAAATGCTGCCAATTCATGCCACCACTTAAATAGAC 136
 DB 152 GCGCAAGAAGAGCTCTGAGTAAGATGGGTCCAAAGCGATGGTCCGCTGCTGATG 211
 QY 137 GACACAAGTATGAGGTGCTGATGAGCTCTACAGATGACCGAGATACCCAAAC 196
 DB 212 GACACCAGAGTGAAGGTGCTGATGAGCTGTACCGGCCACCGAGATTCACGCGAGC 271
 QY 197 AAGAAGAGGACAGAGATCATCAGAACTTCATCAAGACATCAAGCTGAGCGCAT 256
 DB 272 CGCAAGAGGCGCCAGAAATGCTCAAGAACTGTGAAGGTGCGCTGAAAGCTGGAAG 331
 QY 257 CTTATAGAAATTAATGATTATTAATCAAGATGAGCTAGCATTTGAGAGAAATTAAGA 316
 DB 332 CTGCTGCGTGGGAGCAAGCTGGCGGTGAGAGCTGCGCTGCGCGCTTCCGCCAC 391
 QY 317 AAGTTCAATCAGCTGCTATGACCGTGTGATTTTCATCAGGTGATTAATCTTTGAC 376
 DB 392 CGGCGCGCTGCTGCTGATGAGCGCGCTCAGCTTCCACAGGTGAGCTTCACTTGCAC 451
 QY 377 CGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGTGACCAATCATTCAG 436
 DB 452 CGGCGGTGTGAGCGCGCGGCTGCTGAGTCCCGGACCTGTGACCAAGCGCGTGGT 511
 QY 437 CGCAGCTCACTGCGCAAGTCACATGAGCGGGTAAATATATGCTCTTATCATTTTTCAG 496
 DB 512 CCGCAGCTCAAGCGCAAGTCCACGCGCGCATCAACAGATGTGCGCACTGAGCGAG 571
 QY 497 TGTGAATTTTGGCTGCTGTATTAATCTTTTGGGAATTTTAAACCCCATTAACAAA 556
 DB 572 TCGGACTTCTGCTGCTGCTTCAAGCGCGCGCGGACCTTACCGCTCCACCTGCGCAG 631
 QY 557 CTATGATGATGATTAACAATAATGTTGATGAGAGAGAAATATGA 601
 DB 632 ATCTGCGAGGCGCTGGCGCGGATGCTGACGAGGAGGAGCGCTTGA 676

RESULT 24

AD086751
ID AD086751 standard; cDNA; 2186 BP.

XX AD086751;

XX 07-OCT-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #3626.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

XX cancer; cell proliferative disorder; gene; ss.

XX Homo sapiens.

XX WO2004060270-A2.

XX 22-JUL-2004.

XX 15-OCT-2003; 2003WO-US029126.

XX 18-OCT-2002; 2002US-0418988P.

XX (GENENTECH) GENENTECH INC.

XX (WU) WU T D.

XX (ZHOU) ZHOU Y.

XX Wu TD, Zhou Y;

XX WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,

XX preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 3626; 5504bp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identity to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth therapeutically effect of the above protein; (13) a method of cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing

CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX SQ Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;

Query Match 34.0%; Score 204.2; DB 12; Length 2186;

Best Local Similarity 59.3%; Pred. No. 1.3e-49;

Matches 347; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 17 CCGTCGCGGACTCCATGCGCCACAGATGCTTTAAATTCAAAACCTGCGCTTACG 76

Db 92 CGACTCTGCTTCAGAGATGCGCCATGACCTTCACGACCAAGAGCTGCTGCGAG 151

QY 77 GCACAAAAGATCTGGTAAATGTGTCCAAATCCATGCCACCACTTAATAGAC 136

Db 152 GCGCAGAAAGACTCTGATGATGCGTCACAGGACGATGCGCGCTGCTGAT 211

QY 137 GACACAAAGTATGAGGTGCTGATGAGCTCTACAGATGACCGAGATACACCAAAAC 196

Db 212 GACACCAAGATGAGGTGCTGATGAGCTCTACCGCGCCACCGAGGATTCACGGCGAGC 271

QY 197 AAGAAAGGACAGAGATCATCAAGAACTCATCAAGACAGTATCAAGCTGCCATT 256

Db 272 CGCAAGGAGGCCACAGAGATGCTCAAGAACTGCTCAAGTGGCCCTGAAAGCTGGAGACTG 331

QY 257 CTTTATAGGAATATACAGTTTATACAGTATGAGTATGAGAGAAATTAAGAG 316

Db 332 CTGCTGCGGGGACCAAGTGGCGCGGTGAGAGCTGCGCTGCTGCGCGCTTCCGCCAC 391

QY 317 AAGATTCAATCAGCTTGTATGACCGTGTGCTATGCTTTCATCAGTGGATTAATACCTTTGAC 376

Db 392 CGGGGCGCGCTGCTGCGCCATGACGCGCGCTTCCACAGTGAAGCTTCCGCTTCCAGC 451

QY 377 CGGAATGTGTTATCCAGCGCTGTTAAATGATACAGAGATCTGCACTCAATCTACG 436

Db 452 CGCGCGTGTGCTGCGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCG 511

QY 437 GCGCACCTCACGCGCAAGTACATGACGCGGTAAATGCTTGTATGATTTTACGAT 496

Db 512 CCCACCTGACCGCCAGATGCCAGCGCGCATCAACACGATGCTGCGCTGCGCTGCGCGC 571

QY 497 TGTGAATTTTGGCTGCGCTTGTATATCTTTTGGGAAATTTTAAACCCGACTTACAAA 556

Db 572 TCGCATTCCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631

QY 557 CTATGTGATGTATCAACAAAATGTTGATGAAGAGAAATATGA 601

Db 632 ATCTCGAGAGGCTGCGCGCGATGCTGACGAGGCGCAGCTCTGA 676

RESULT 25

AD084271
ID AD084271 standard; cDNA; 2186 BP.

XX AD084271;

XX 07-OCT-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #1085.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

XX cancer; cell proliferative disorder; gene; ss.

XX Homo sapiens.

XX WO2004060270-A2.

XX 22-JUL-2004.

XX 15-OCT-2003; 2003WO-US029126.

XX 18-OCT-2002; 2002US-0418988P.

XX (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 PI WU TD, Zhou Y;
 XX WPI, 2004-534300/51.
 DR
 XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX
 XX Claim 1; SEQ ID NO 1085; 5504bp; English.
 CC The present invention describes an isolated tumour-associated antigenic
 CC target (TAR) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAR)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAR binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAR sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAR cDNA sequence from the present invention.
 XX
 XX Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;
 SO
 Query Match 34.0%; Score 204.2; DB 13; Length 2186;
 Best Local Similarity 59.3%; Pred. No. 1.3e-49;
 Matches 347; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
 QY 17 CGGTGCGGCACTCTCCGATGAGCCACAGATGCTTTAATCCAAAAACCTGGCCGTTCCAG 76
 DB 92 CGGACTCTGCTTGAGAGTAGGCGCAGACACCTTCGACCAAGAGCTGGCTTCGAG 151
 QY 77 GCACAAAAGAGATCTTGGTAAATGNTGCCAAATCCATCCACCACTTAATAGAC 136
 DB 152 GCGCAGAAAGAGCTCCGATGAAGATGCGTCCAGCAATGAGCGCGTGTGGAT 211
 QY 137 GACACAGTAGTAGGTGCTGATGAGCTTACAGAGTGCAGGAGTAGTACCCCAAC 196
 DB 212 GACACCAAGCAGTAGGTGCTGATGAGCTTACCGGCCCAAGGAGTTCAAGCGCAGC 271
 QY 197 AAGAGGAGGAGAGATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCATT 256

DB 272 CGCAAGAGAGGCCCAAGAAAGATGCTCAAGAACCTGTGTAAGTGCCCTGAAGCTGGACATG 331
 QY 257 CTTTATAGAAATATATCATGTTTAAATCAAGATGAGCTAGCATTTGAGAAATTTAGAG 316
 DB 332 CTGCTGCTGGGAGCAGCTGGAGCGGTGAGAGAGCTGGCTGCGCGCTTCGCGCAC 391
 QY 317 AAAGTTATCATGCTTGTGATAGACCGGTGTAGATTTTCCATAGGTGATTTACTTTGAC 376
 DB 392 CGGCGCGCTGCTGGCGCATAGACGGCCGTAGCTTCACACAGGTGGACTTCACCTTGAC 451
 QY 377 CGGAATGTGTATTCACAGGCTGTTAAATGAATGACAGAGATGTGCACCAATGATTCAG 436
 DB 452 CGGCGCTGCTGGCGCTGGCGCGCTGCTGATGCTCCGACCTGCTGCACCGCGCTGGT 511
 QY 437 CGCAGCTCTACTGCCAAGTACATGACGCGGTTAATTAATGCTTTGATCATTTTTCAGAT 496
 DB 512 CCCAGCTGACCGCCCAAGTCCACGCGCGCATCAACACAGTGTTCGCGCACCTGACCGAC 571
 QY 497 TGTGAATTTTGGCTGCTGTATTAATCTTTTGGGAATTTTAAACCCACTTACAAA 556
 DB 572 TGGACTTCTGTGCTGTGCTCTACGCGCCCGCAGCCCTTACCGCTTCGCGCAGG 631
 QY 557 CTATGAGATGATATCAAAAATGTTGATGATGATGATGATGATGATGATGATGATGAT 601
 DB 632 ATCTGAGAGGCTCTGGCGCGATGCTGAGACGAGGACCTTCTGA 676
 RESULT 26
 ABS78742
 ID ABS78742 standard; DNA; 619 BP.
 AC ABS78742;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE DNA encoding human NOX17b protein.
 XX
 KW Human; NOX1; human disease; NOX1-associated disorder; cancer; addiction;
 KW Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
 KW tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
 KW Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
 KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;
 KW obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
 KW infertility; inflammatory bowel disease; atherosclerosis; hypertension;
 KW scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
 KW asthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
 KW bacterial infection; parasitic infection; graft-versus-host disease;
 KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
 KW angiogenesis; gene; de.
 XX
 OS Homo sapiens.
 OS
 PN WO200272770-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US007283.
 XX
 PR 08-MAR-2001; 2001US-0274281P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 20-MAR-2001; 2001US-0277232P.
 PR 20-MAR-2001; 2001US-0277237P.
 PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.

CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention
 XX
 XX Sequence 1113 BP; 282 A; 311 C; 304 G; 216 T; 0 U; 0 Other;

Query Match 30.3%; Score 182; DB 6; Length 1113;
 Best Local Similarity 58.7%; Pred. No. 3.6e-43;

Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGGCCACAGATGCTTTAATCCAAAAACCTGGCCGTTCCAGGACAAAGAAGA 89
 DB 52 CTCAGAGACCCATGAGAGTCTTCAGCTCAAGAGCCTGGACATCGACAGAGAGAAGC 111
 QY 90 TCTTGGGTAAATGCTGTCCAAATCCATGCCACCATTTAATAGACAGACAAGTAGTG 149
 DB 112 TACTGAGTAGATGAGCGCGGTGCTGTGGCTCATCTTCATAGATGAGACAAGCAGTG 171
 QY 150 AGGTGCTGATGAGCTCTACAGAGTGCAGAGAGATACCCCAAAACAAAGAGGACG 209
 DB 172 AGGTGCTGATGAGCTCTACAGAGTGCAGAGAGATACCCCAAAACAAAGAGGACG 221
 QY 210 AGAAGATCATCAAGAACCTCATCAAGACAGTCAAGTCAAGTGCATTTATAGAAATA 269
 DB 232 AGCGCGTATCAAGAGACCTATCAAGTGCATCAAGTGCCTGCTGCTCACCAGCATG 291
 QY 270 ATCAGTTTAAATCAAGATGAGCTAGCATGTGATGAGAAATTTAAGAAAGATTATCAGC 329
 DB 292 GCTCTTTGGCCCCAGAGAGCTGACCTGCTACCCGCTTTCCAGAAAGCTCGGAGG 351
 QY 330 TTGCTATACCGGTGCTGATTTCCATCAGTGTGATTAATCTTTGACCGGAATGTGTAT 389
 DB 352 GTGCCATGAGGAGCACTTAGCTTTGGTGAAGTAACCTTCAGAGGCTGCTGCTTGG 411
 QY 390 CCAGGCTGTTAATGAATGAGAGAGATGCTGCACCAATCATTCAGCGCCACTCAGT 449
 DB 412 CTGGCTGTGACCGAGAGTCCGGGAGTGTCTCTAAGTTGGTGAACACCACTTACGC 471
 QY 450 CCAAGTCATGAGACGGGTTAATATGCTTTGATCATATTTTTCAGATTGTGAATTTTGG 509
 DB 472 CCAAGTCATGAGACGGGTTAATATGCTTTGATCATATTTTTCAGATTGTGAATTTTGG 531
 QY 510 CTGGCTGTGATATCTTTTGGGAATTTTAAACCCACTTACAAAACATGTGATGTGA 569
 DB 532 CGGCGCTTATGAGGCTCT-----GACTTCACTCAGACACCTTGGCAAGATCTGTAGCGAC 585
 QY 570 TCACAAAATGTTGATGAGAGACATATGA 601
 DB 586 TCAGGAAGCTCTAGACGAAGGAGAGCTTGA 617

RESULT 28
 AAD38699
 ID AAD38699 standard; cDNA; 1165 BP.
 XX
 XX AAD38699;
 DT 23-SEP-2002 (first entry)
 XX
 XX Human LP237 secreted protein encoding cDNA.
 DB
 XX Human; secreted protein; atherosclerosis; Alzheimer's disease; LP237;
 KW diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
 KW rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
 KW reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
 KW gene therapy; neoplasm; transgenic; psoriasis; ischemia; carcinoma;
 KW gene; ss.

XX
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 100..654
 FT /tag= a
 FT /product= "human LP237 secreted protein"
 FT sig_peptide 100..180
 FT /tag= b
 FT mat_peptide 181..651
 FT /tag= c
 FT /product= "Mature human LP237 secreted protein"
 XX
 PN WO200226801-A2.
 XX
 XX 04-APR-2002.
 PD
 XX
 XX 14-SEP-2001; 2001WO-US026026.
 PF
 XX 28-SEP-2000; 2000US-0236088P.
 PR
 XX (ELIL) LILLY & CO ELI.
 PA
 XX
 XX Su EW, Wang H;
 PI
 XX WPI; 2002-471259/50.
 DR P-PSDB; AAE23984.
 DR
 XX
 PT Novel proteins and polynucleotides of secreted proteins useful for
 PT treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
 PT diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury.
 PT
 XX
 XX Claim 1; Page 140-141; 145pp; English.
 PS
 XX
 XX The invention relates to human secreted polypeptides designated LP095,
 CC LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic acid
 CC molecules encoding such polypeptides. Novel secreted proteins of the
 CC invention are used for treating diseases such as atherosclerosis,
 CC Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
 CC arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
 CC combined immunodeficiency, ischemia, carcinoma, haemolytic anaemia,
 CC reperfusion injury, neoplasms and cancer especially liver cancer. They
 CC are also used for wound healing. Polynucleotides of the invention can be
 CC used to generate transgenic animals or knock out animals, which in turn,
 CC are useful for use in the development and screening of therapeutically useful
 CC reagents for use in the treatment of diseases associated with LP
 CC polypeptide associated activity. They are also used in gene therapy. The
 CC present sequence is human LP237 secreted protein encoding cDNA
 CC
 XX
 SQ Sequence 1165 BP; 305 A; 319 C; 312 G; 229 T; 0 U; 0 Other;

Query Match 30.3%; Score 182; DB 6; Length 1165;
 Best Local Similarity 58.7%; Pred. No. 3.7e-43;
 Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGGCCACAGATGCTTTAATCCAAAAACCTGGCCGTTCCAGGACAAAGAAGA 89
 DB 89 CTCAGAGACCCATGAGAGTCTTCAGCTCAAGAGCCTGGACATCGACAGAGAGAAGC 148
 QY 90 TCTTGGGTAAATGCTGTCCAAATCCATGCCACCATTTAATAGACAGACAAGTAGTG 149
 DB 149 TACTGAGTAGATGAGCGCGGTGCTGTGGCTCATCTTCATAGATGAGACAAGCAGTG 208
 QY 150 AGGTGCTGATGAGCTCTACAGAGTGCAGAGAGATACCCCAAAACAAAGAGGACG 209
 DB 209 AGGTGCTGATGAGCTCTACAGAGTGCAGAGAGATACCCCAAAACAAAGAGGACG 268
 QY 210 AGAAGATCATCAAGAACCTCATCAAGACAGTCAAGTCAAGTGCATTTATAGAAATA 269
 DB 269 AGCGCGTATCAAGAGACCTATCAAGTGCATCAAGTGCCTGCTGCTCACCAGCATG 328
 QY 270 ATCAGTTTAAATCAAGATGAGCTAGCATGTGATGAGAAATTTAAGAAAGATTATCAGC 329

Db 329 GCTCCTTTGGCCCCAGTGGAGCTGGCCCTGGCTTTCGCCAGAAAGCTGGCAGG 388
 QY 330 TTGCTATGACCGGTGTCAGTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTGTAT 389
 Db 389 GTGCCATGACGGACTAGCTTTGGTGAAGTAACTTCACTTCAGAGCTGCTGTCTGG 448
 QY 390 CCAGGCTGTTAAATGATGACAGAGATGCTGCACCAATCATTTCCAGCCCTCACTG 449
 Db 449 CTGGCTGTGACCCGATGCGGGGATGTGCTGTAAAGTTGGTGAACACCACTCAAGC 508
 QY 450 CCAAGTCACATGACGGGTTAAATGCTTTGATCATATTTTCAATTTGTAATTTTGG 509
 Db 509 CCAAGTCACATGACGGGCTATCCGACATGCTTTGATCATCTCTGACCCAGGCTGCTCA 568
 QY 510 CTGCTCTGTATATCCTTTTGGGAATTTTAAACCCACTTACAAAACATATGTGATGTA 569
 Db 569 CGGCTCTATATGGGCTT-----GACTTCACTAGACACCTTGGCAAGATCTGTGAGCAG 622
 QY 570 TCACAAAAATGTTGATGAAGAACATATGA 601
 Db 623 TCAGGAAGCTGCTAGACGAAGGAAGCTCTGA 654
 RESULT 29
 ID ABZ11655 standard; cDNA; 1175 BP.
 AC ABZ11655;
 DT 20-JUN-2003 (first entry)
 DE Human polynucleotide SEQ ID NO 537.
 XX
 KM Human; genome mapping; gene therapy; food supplement; virus; fungus;
 cell-proliferative disorder; neurodegenerative disease; bacterial;
 Parkinson's disease; Alzheimer's disease; autoimmune disease;
 multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 arthritis; cytotoxic; immunomodulator; neotropic; neuroprotective;
 anti-inflammation; antidiabetic; immunosuppressive; dermatological;
 haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
 antidiabetic; gene; ss.
 KM
 OS Homo sapiens.
 XX
 PN MO200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002MO-US005095.
 XX
 PR 05-MAR-2001; 2001US-00799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Wehrman T, Wang J, Wang D, Dzmanac RT;
 XX
 DR WPI; 2002-759812/82.
 DR P-PSDB; ABP69438.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.
 XX
 PS Claim 1; SEQ ID NO 537; 1012bp + Sequence Listing; English.
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
 CC ABZ12066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The

CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from Wipo at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1175 BP; 306 A; 324 C; 313 G; 232 T; 0 U; 0 Other;
 Query Match 30.3%; Score 182; DB 6; Length 1175;
 Best Local Similarity 58.7%; Pred. No. 3.7e-43;
 Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;
 QY 30 CTCGATGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTACAGGACAAAAGAGA 89
 Db 102 CTCGAGAACCCATGAGATCCTTCAAGCTCAAAAGCTGGCACTGCAAGAGAAAGAGC 161
 QY 90 TCTTGGTAAATGATGTCCTCAATTCATCGCCACCACTTAATYAGACACAGATAGTG 149
 Db 162 TACTGATGAAGATGGCGGGTCCCTCTGTGGCTCATCTCTCATATGATGAGACACAGATG 221
 QY 150 AGGTGCTGATAGATCTTACAGATGACAGGAGATACCCAAAACAAAGAGGGCAG 209
 Db 222 AGGTGCTGATAGATCTTACAGATGACAGGAGATACCCAAAACAAAGAGGGCAG 281
 QY 210 AGAAGATCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATTTCTTTATAGAAATA 269
 Db 282 AGCGGTATCAAGAGACCTGATCAAAAGTGCATCAAGTGGCTGTGCTGCAACCGAAG 341
 QY 270 ATCAAGTTAATCAAGATGAGTACATGATGATGAGAGAAATTAAAGAAAGTTCAATCAGC 329
 Db 342 GCTCCTTTGGCCCCAGTGGAGCTGGCCCTGGCTAACCCGCTTTGGCCAGAAAGCTGGCCAGG 401
 QY 330 TTGCTATGACCGGTGTCAGTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTGTAT 389
 Db 402 GTGCCATGACGGCACTTACCTTTGGTGAAGTCACTTCAAGAGCTGCTGTCTGG 461
 QY 390 CCAGGCTGTTAATGATGACAGAGATGCTGCACCAATCATTTGAGCCGCACTCACTG 449
 Db 462 CTGGCTGTGACCCGATGCGGGGATGTGCTGTAAAGTTGGTGAACACCACTCAAGC 521
 QY 450 CCAAGTCACATGACGGGTTAAATGCTTTGATCATATTTTCAATTTGTAATTTTGG 509
 Db 522 CCAAGTCACATGACGGGCTATCCGACATGCTTTGATCATCTCTGACCCAGGCTGCTCA 581
 QY 510 CTGCTCTGTATATCCTTTTGGGAATTTTAAACCCACTTACAAAACATATGTGATGTA 569
 Db 582 CGGCTCTATATGGGCTT-----GACTTCACTAGACACCTTGGCAAGATCTGTGAGCAGC 635
 QY 570 TCACAAAAATGTTGATGAAGAACATATGA 601
 Db 636 TCAGGAAGCTGCTAGACGAAGGAAGCTCTGA 667
 RESULT 30
 ID ADM44173 standard; cDNA; 1175 BP.
 AC ADM44173;
 DT 03-JUN-2004 (first entry)
 DE Novel human arginine-rich protein cDNA #537.
 XX
 KM ss; gene; human; arginine-rich protein; cancer; inflammation;
 KM genetic disorder.
 XX
 OS Homo sapiens.

```

XX  US2004053250-A1.
PN
PD  18-MAR-2004.
PF  21-NOV-2002; 2002US-00302172.
PR  05-MAR-2001; 2001US-00799451.
PR  05-MAR-2002; 2002WO-US005095.
PR  20-AUG-2002; 2002US-00225251.
XX
XX  (TANG/) TANG Y T.
PA  (XUEA/) XUE A.
PA  (DRMA/) DRMANAC R T.
XX
XX  Tang YT, Xue A, Drmanac RT;
PI
XX  WPI; 2004-238579/22.
XX
XX  New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT  useful for diagnosing and/or treating conditions associated with aberrant
PT  activity of the arginine-rich polypeptides, such as cancer and
PT  inflammation.
XX
XX  Disclosure; SEQ ID NO 537; 51pp; English.
XX
XX  The invention relates to an isolated polynucleotide. The methods and
CC  compositions of the present invention are useful for the diagnosis and/or
CC  treatment of diseases or conditions associated with aberrant expression
CC  or activity of the arginine-rich protein-like polypeptides, such as
CC  cancer and inflammation. They can also be used in forensics, gene
CC  mapping, identification of mutations responsible for genetic disorders,
CC  and in assessing biodiversity. The present sequence represents a novel
CC  human arginine-rich protein cDNA.
XX
XX  Sequence 1175 BP; 306 A; 324 C; 313 G; 232 T; 0 U; 0 Other;
SQ
Query Match 30.3%; Score 182; DB 12; Length 1175;
Best Local Similarity 58.7%; Pred. No. 3,7e-43;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;
QY 30 CTCGATGGCAGACAGATGCTTTAATTCAGAAACCTGGCCCTTCAGGCACAAAGAGA 89
DB 102 CTCAGAGACCCATGAGTCTTCTTACCTCAAGAGCCTGGACCTGCAGAGAGAGAAG 161
QY 90 TCTTGGGTAAATGATGTCCTCAATTCATCGCCACCACTTAATAGACACAAAGTAGT 149
DB 162 TACTGATGATGAGTGGGGGCTGCTGTGCTCATCTCTTCATAGATGACACAGCAGT 221
QY 150 AGGTGCTGATGAGTCTTACAGAGTACAGGAGTACACCCAAACAAAGAGAGGCGAG 209
DB 222 AGGTGCTGATGAGTCTTACAGAGTACAGGAGTACAGGAGGCGGCGCCAGGCC 281
QY 210 AGAAGATCATCAAGACCTCATCAAGACATCATCAAGCTGGCATTTCTTATAGAAAT 269
DB 282 AGGCGCTGATCAAGACCTCATCAAGTGCATCAAGGCGTGTGCTGCACCGGAATG 341
QY 270 ATCAGTTTAAATCAAGATGAGTACATGATGAGAAATTAAGAAGAAATTCATCAG 329
DB 342 GCTCCTTTGGCCCCAGTAGCTGGCCCTGCTACCCGCTTCCCGAAGAGCTCGGCG 401
QY 330 TTGCTATGACCGTGTGCTGATTCATAGTGTATATACCTTTGACCGGATGTGTAT 389
DB 402 GTGCCATGACGCGCATTAAGCTTTGAGAGGTATCACTTCGAGGCTGTGTCTGG 461
QY 390 CAGGCGTGAATGAATGATGAGAGAGATGTGACCAAAATCATTCAGCGCCACTCAG 449
DB 462 CTGGCTGTCTGACCGAGTGGCGGAGTGTGCTGAGGTGTGGAACACCACTCAGCG 521
QY 450 CCAAGTCACATGACGCGGTAAATATATGCTTTGATCATTTTTCAGATTGTGAATTTTGG 509
DB 522 CCAAGTCACATGACGCGCGGATCCGCGAGGTGTTGATCATCTTCTGACCCAGGCTCTGCTCA 581

```

```

QY 510 CTGGCTGTGATATATCCTTTGGGAATTTAAACCCGACTTACAAAGACTATGATGATGTA 569
DB 582 CGGCTCTTATGGGCTT-----GACTTCATCTAGACCTTGGCAAGATCTGTAGCGAC 635
QY 570 TCACAAAAATGTTGATGAGAGACATATGA 601
DB 636 TCAGAGAGCTGCTAGACGAAAGGAGAGACTCTGA 667

RESULT 31
AAF59594
ID AAF59594 standard; cDNA; 1268 BP.
XX
AC AAF59594;
XX
XX 24-APR-2001 (first entry)
XX
XX Human cell cycle and proliferation protein CCYPR-5 cDNA, SEQ ID NO:59.
XX
XX Cell cycle and proliferation protein; CCYPR; human; agonist; antagonist;
KW gene therapy; detection; gene therapy; transgenic animal disease model;
KW immune disorder; developmental disorder; cell signalling disorder;
KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection; ss.
XX
XX Homo sapiens.
XX
XX MO200107471-A2.
XX
XX 01-FEB-2001.
XX
XX 21-JUL-2000; 2000WO-US019948.
XX
XX 21-JUL-1999; 99US-0145075P.
XX 08-SEP-1999; 99US-0153129P.
XX 10-NOV-1999; 99US-0164647P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DM, Baughn MR, Patterson C, Shah P;
DR P-PSDB; AAB60457.
XX
XX WPI; 2001-112727/12.
XX
XX Human cell cycle and proliferation proteins and polynucleotides are used
PT to treat, diagnose and prevent immune, developmental and cell signaling
PT disorders and cell proliferative disorders including cancer.
XX
XX Claim 5; Page 171; 205pp; English.
XX
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human cell cycle
CC and proliferation proteins (CCYPR), AAB60453-AAB60506. CCYPR and agonists
CC of CCYPR are used to treat diseases or conditions associated with
CC decreased expression of functional CCYPR, while CCYPR antagonists are
CC used to treat diseases or conditions associated with overexpression of
CC functional CCYPR. Monoclonal or polyclonal antibodies to CCYPR may be
CC used in enzyme-linked immunosorbent assays (ELISA) or radioimmunoassays to
CC detect CCYPR. CCYPR itself may be used to detect compounds e.g.,
CC antibodies, oligonucleotides and proteins (receptors) that specifically
CC bind to CCYPR, and in drug screening methods to identify compounds that
CC modulate the activity of CCYPR. CCYPR nucleotides can be used to generate
CC transgenic animal models of human disease, and can be used in gene
CC therapy in target cells with genetic abnormalities with respect to the
CC expression of CCYPR for the treatment or prevention of a disorder
CC associated with CCYPR. Diseases which can be diagnosed, treated and
CC prevented using CCYPR proteins, nucleic acids, agonists or antagonists
CC include immune, developmental and cell signalling disorders, and cell
CC proliferative disorders including cancer. Specific examples of these
CC disorders include anaemia, epilepsy, arteriosclerosis, asthma, cancer,
CC allergies, diabetes mellitus, disorders of the menstrual cycle and
CC infections caused by bacteria

```

XX Sequence 1268 BP; 343 A; 345 C; 350 G; 230 T; 0 U; 0 Other;
 SQ Query Match 30.3%; Score 182; DB 4; Length 1268;
 Best Local Similarity 58.7%; Pred. No. 3.9e-43;
 Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCCTGAGGACACAGATGCTTTAATCCAAAACCTGGCCGTTGAGGACAAAAGAAGA 89
 DB 92 CTCAGAGACCCATGAGTCTTTCAGCTCAAGAGCTGGCACTGCAAGCAGAGAAAGC 151

QY 90 TCTTGGGTAAATGCTGCCAATCCATGCGCCACCACTTAATAGACGACAAAGTAGTG 149
 DB 152 TACTAGTAAAGATGGGGGTGCTGTGTGGCTATCTCTTCATAGATGAGACAGAGTG 211

QY 150 AGGTGCTGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAGAGAGCGAC 209
 DB 212 AGGTGCTAATGAGCTCTACCGGTGTCCAAAGAGTACACGACACCGCCCGCCGCGCC 271

QY 210 AGAAGATCATCAAGACCTCATCAAGACAGTCAAGTCCATTCATTTAATAGATA 269
 DB 272 AGCGCTGATCAAGAGCTGATCAAGTGCATCAAGGTGCTGTGCTGACCGCAATG 331

QY 270 ATCAGTTTAATCAAGATGAGTACGATTGATGAGAAATTTAAGAAAGTTTCATCAGC 329
 DB 332 GCTCCCTTGGCCCACTGAGCTGGCCCTGGCTACCCGCTTTGGCCAGAAAGCTGGCGACG 391

QY 330 TTGCTATGACCGTGTGCTGCTTCATCAGTGGATTAATACCTTTGACCGGAATGTGTAT 389
 DB 392 GTGGCCATGACCGGACTTAGCTTTGGAGGTATACCTTCAGAGCTGTCTGTG 451

QY 390 CCAAGGCTGTTAATGATGATGAGAGATGCTGCACCAAAATCATTCAGCGCCACCTCAGT 449
 DB 452 CTGGCTGTGCTGACCGGCTGCGGAGTGTGCTTAAGTGTGAGAACACACCTCAACG 511

QY 450 CCAAGTACATGACCGGCTTAAATATGCTTTTGCATTTTTCAGATTGTAATTTTGG 509
 DB 512 CCAAGTACATGACCGGCTTAAATATGCTTTTGCATTTTTCAGATTGTAATTTTGG 571

QY 510 CTGGCTGTGATTAATCCTTTTGGAAATTTTAAACCCCACTTCAAAAACATGATGATGTA 569
 DB 572 CGGCTCTTATGAGGCT-----GACTTCACTGACGACCTTGGCAAGATCTGTGAGCGAC 625

QY 570 TCACAAAAATGTTGATGAAGAGACATATGA 601
 DB 626 TCAGGAAGCTGCTAAGCAAGGAGAGCTCTGA 657

RESULT 32
 ACH76016
 ID ACH76016 standard; DNA; 500 BP.
 XX ACH76016;
 XX 29-JUL-2004 (first entry)
 XX Human genome derived single exon probe #9211.
 XX Human; probe; ss; gene expression; single exon probe; microarray;
 XX alternative splicing event; genomic alteration.
 XX Homo sapiens.
 XX US2003194704-A1.
 XX 16-OCT-2003.
 XX 03-APR-2002; 2002US-00029386.
 XX 03-APR-2002; 2002US-00029386.
 XX (PENN/) PENN S G.
 XX (RANK/) RANK D R.
 XX PA

PA (HANZ/) HANZEL D K.
 XX Penn SG, Rank DR, Hanzel DK;
 XX WPI; 2004-119264/12.
 XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX Claim 15; SEQ ID NO 9211; 80bp; English.
 XX The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC a method of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX Sequence 500 BP; 164 A; 71 C; 107 G; 158 T; 0 U; 0 Other;
 SQ Query Match 28.0%; Score 168.4; DB 12; Length 500;
 Best Local Similarity 99.4%; Pred. No. 2.6e-39;
 Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 432 TTCAGGCACTCTGCTGCTGCTGATTAATCCTTTGGAATTTTAAACCCCACTTAC 491
 DB 1 TTCAGGCACTCTGCTGCTGCTGATTAATCCTTTGGAATTTTAAACCCCACTTAC 60

QY 492 CAGATTGTAATTTTGGCTGCTGCTGATTAATCCTTTGGAATTTTAAACCCCACTTAC 551
 DB 61 CAGATTGTAATTTTGGCTGCTGCTGATTAATCCTTTGGAATTTTAAACCCCACTTAC 120

QY 552 AAAAATATGATGATGATCAACAAAATGTTGATGAAGAGAAACATATGA 601
 DB 121 AAAAATATGATGATGATCAACAAAATGTTGATGAAGAGAAACATATGA 170

RESULT 33
 AAC10352
 ID AAC10352 standard; cDNA; 340 BP.
 XX AAC10352;
 XX

XX 06-OCT-2000 (first entry)
DT Human secreted protein 5' EST, SEQ ID NO: 14427.
XX
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX Gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX BP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX
XX (GSEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclet A, Giordano J;
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX
XX Claim 1; SEQ ID NO 14427; 71bp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNA encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 340 BP; 92 A; 79 C; 85 G; 79 T; 0 U; 5 Other;
Query Match 26.5%; Score 159.2; DB 3; Length 340;
Best Local Similarity 98.8%; Pred. No. 1.2e-36;
Matches 158; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 45 ATGCTCTTAATTCACAAAACCTGGCGCTTCAGGCACACAAAAGAGATCTTGGTAAATG 104
DB 181 ATGCTCTTAATTCACAAAACCTGGCGCTTCAGGCACACAAAAGAGATCTTGGTAAATG 240
QY 105 TGTCCAAATTCATCGCCACCACTTAATAGACACACAACTAGTAGAGTGTGATGAGC 164
DB 241 TGTCCAAATTCATCGCCACCACTTAATAGACACACAACTAGTAGAGTGTGATGAGC 300
QY 165 TCTACAGATGACACAGGAGTAGACCCCAAAACAAAGAGA 204
DB 301 TCTACAGATGACACAGGAGTAGACCCCAAAACAAAGAGA 340
RESULT 34
ABL07183
ID ABL07183 standard; cDNA; 829 BP.
XX ABL07183;
AC
XX 26-MAR-2002 (first entry)
DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 16031.
XX

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX
XX P-PDB; ABB63080.
XX
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX
XX Claim 1; SEQ ID NO 16031; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AAB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO.int/pub/pat/published_pat_sequences
XX
SQ Sequence 829 BP; 229 A; 208 C; 210 G; 182 T; 0 U; 0 Other;
Query Match 24.3%; Score 146.2; DB 4; Length 829;
Best Local Similarity 57.6%; Pred. No. 1.3e-32;
Matches 262; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
QY 35 ATGGCAGACAGATGTCTTAAATTCACAAAACCTGGCGCTTCAGGCACACAAAAGAGATCTTG 94
DB 263 ATGGCAGACAGATGTCTTAAATTCAGATTCGAGCTGGGCGCAGAAAGATCTTC 322
QY 95 GGTAAATGTGTCCAAATTCATCGCCACCACTTAATAGACACACAACTAGTAGAGTG 154
DB 323 TCACGAGATGCCACAAAACATCGCCACCACTTAATAGAGACGAGCGGTGCGTG 382
QY 155 CTGATGAGCTCTACAGAGTACACAGGAGTACACCCAAACAAAGAGAGGACGAGAG 214
DB 383 CTGACCAACCTCTACAGGCTCTGCAAGATGACACAGGAGAAAGGCGGAGAGAG 442
QY 215 ATCATCAAGACCTCATCAAGACAGTCATCAAGCTGCGCATCTTATAGAAATATCG 274
DB 443 CTGATCAAGACCTCATCAAGATGATCAAGATGCTGCTCCACCGAACAATCG 502
QY 275 TTTAATCAAGATGAGCTAGATGAGAAATTTAAGAAAGATTATCATGAGTGTGCT 334
DB 503 TTCAGGAGACGAGAGAGTGCAGAAAGGCGAGCTTTTAAGAAAGTTCAAAACAGCGAG 562
QY 335 ATGACCGTGTGATGATTCATCAGGTGATTTATCTTGAACCGGAATGTATTATCAG 394
DB 563 CTGTCCATCATATCAATTTACAGAGTGAAGCTTCAGAGTTCAGCTGCGGACGAG 622
QY 395 CTGTTAATGATGACGAGAGATGCTGCACCAATCATTCAGGCGACCTGACGAG 454
DB 623 TCATATGCGGAATCGCAAGTGGCGCTGAAGTGTGATGCTGACCGCATCTCACAAAG 682

Query Match 21.0%; Score 126.4; DB 3; Length 252;
Best Local Similarity 98.6%; Pred. No. 5.4e-27;
Matches 138; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTTCAAGCTCCGCGCGCGTGC-CGACTCTCGATGGCCACAGATGCTTTAATTCCA 59
Db 113 CTTCAAGCTCCGCGCGCGTGC-CGACTCTCGATGGCCACAGATGCTTTAATTCCA 112

QY 60 AAAACCTGGCGCTTCAGGCAAAAAGATCTTGGGTAAATGGTGTCCAAATCCATCG 119
Db 173 AAAACCTGGCGCTTCAGGCAAAAAGATCTTGGGTAAATGGTGTCCAAATCCATCG 232

QY 120 CCACCACTTAATAGACGAC 139
Db 233 CCACCACTTAATAGACGAC 252

RESULT 37
AAK56622
ID AAK56622 standard; cDNA; 441 BP.
XX
AC AAK56622;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1682.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220863P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241877P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246174P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246539P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.

QY 39 CCACGAGATGTCCTTTAATTCGAAAAACCTGGCCGTTACAGCACAAAAAGAAATCTTGGGTA 98
 Db 78 CCATGGAGTCTTTCAGCTCCAAAGACCTTGGCACTGCACACAGAAAGAAAGAACTACTAGAGTA 137
 QY 99 AATGTGTCCAAATTCATTCGCCACACACCTTTATATGACACCAAGTAAGTAGTGCTGG 158
 Db 138 AGATGGCGGTCGCTCTGTGGCTCATCTCTTATGATGATGACCAAGCATGTGGTGCTG 197

Seq	Sequence	4406 BP	1178 A	1021 C	1027 G	1180 T	0 U	0 Other
Query Match	19.3%	Score 116	DB 4	Length 4406				
Best Local Similarity	60.4%	Pred. No. 2.8e-23						
Matches 191	Conservative	0	Mismatches 125	Indels 0	Gaps 0			

Db 1650 TCCACCCCTCCATCCGCTTACTCATGGCGACAAATGCTTCAAGTCGACGATATCG 1591
Qy 69 CCGTTCAGGCACAAAAGATCTTGGTAAATGCTGTCCTCAATTCATGCGCACCT 128
Db 1590 GGCTGCGGGCGCAAAAGATCTCTTCACGATGCGCACAAAAGATGCGCAAGACT 1531
Qy 129 TAATAGACGACAAAGTAGTAGAGGTGCTGATGAGCTCTTACAGAGTACCGAGGAGATACA 188
Db 1530 TCATCATGTCGACACGCGCTGCTGTCGACAACTCTTACAGGCTCTGCAAGATGACACA 1471
Qy 189 CCGCAAAACAAGAGAGGACAGAGATCATCAAGAACTTCATCAAGACATCATCAAGC 248
Db 1470 CCGGGAACAAGGCGCAAGGAGAGATCATCAAGAACTTCATCAAGATGATGATCAAGA 1411
Qy 249 TGGCCATCTTTATAGAAATATATAGTTTAAATCAAGATGAGCTTATGATGAGAAAT 308
Db 1410 TCGGTGTGCTCCACCGGACAAATCATGTCAGCGACGAGAGCTGCAGAGCGGAGCTCT 1351
Qy 309 TTAAGAGAAAGTTCA 324
Db 1350 TTAAGAGAAAGTTCA 1335

RESULT 39

AA192435/c

AA192435 standard; cDNA; 400 BP.

AA192435;

06-NOV-2001 (first entry)

Human polynucleotide SEQ ID NO 12495.

Human; cytokine; cell proliferation; cell differentiation; gene therapy;
vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
tissue growth factor; immunomodulatory; cancer; leukaemia;
nervous system disorders; arthritis; inflammation; ss.

Homo sapiens.

MO200164835-A2.

07-SEP-2001.

26-FEB-2001; 2001MO-US004927.

28-FEB-2000; 2000US-00515126.

18-MAY-2000; 2000US-00577409.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-514838/56.

P-PDB; AA012504.

Isolated nucleic acids and polypeptides, useful for preventing diagnosing

and treating e.g. leukemia, inflammation and immune disorders.

Claim 1; SEQ ID NO 12495; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and
the encoded proteins (AA000010-AA013910) that exhibit activity relating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activity/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format

CC directly from WPI at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 400 BP; 80 A; 122 C; 114 G; 84 T; 0 U; 0 Other;
SQ
Query Match 16.64; Score 99.8; DB 4; Length 400;
Best Local Similarity 68.84; Pred. No. 5.4e-19;
Matches 137; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 23 CCGACTCTCCGATGCGCACAGATGCTTAAATTCAAAAACCTGCGCTTCAGGCACAA 82
Db 199 CTGATGTCCTCCGAGGCGATGACACCTTCAGCACCAAGAGCTGCTGCAAGCGCAG 140
Qy 83 AAGAAGATCTTGGTAAATGCTGTCCTCAATTCATGCGCACCTTAATAGACACACA 142
Db 139 AAGAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 80
Qy 143 AGTAGAGAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 202
Db 79 AGCAGTAGAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20
Qy 203 GAGGACAGAGAGATCATCA 221
Db 19 GAGGCCAGAGAGATGCTCA 1

RESULT 40

AAH04721

AAH04721 standard; cDNA; 714 BP.

AAH04721;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:1556.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-00116126.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

11-JAN-2000; 2000JP-00118776.

02-MAY-2000; 2000JP-00183767.

09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Claim 1; SEQ ID NO 1556; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-
length cDNAs defined in the specification. Where a primer set comprises:
(a) an oligo-dT primer and an oligonucleotide complementary to the
complementary strand of a polynucleotide which comprises one of the 5602
cDNAs defined in the specification; and for the detection and/or
diagnosis of the abnormality of the proteins encoded by the full-length
cDNAs.

Claim 1; SEQ ID NO 1556; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-
length cDNAs defined in the specification. Where a primer set comprises:
(a) an oligo-dT primer and an oligonucleotide complementary to the
complementary strand of a polynucleotide which comprises one of the 5602
cDNAs defined in the specification; and for the detection and/or
diagnosis of the abnormality of the proteins encoded by the full-length
cDNAs.

Claim 1; SEQ ID NO 1556; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-
length cDNAs defined in the specification. Where a primer set comprises:
(a) an oligo-dT primer and an oligonucleotide complementary to the
complementary strand of a polynucleotide which comprises one of the 5602
cDNAs defined in the specification; and for the detection and/or
diagnosis of the abnormality of the proteins encoded by the full-length
cDNAs.

CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX
SQ Sequence 714 BP; 113 A; 246 C; 203 G; 148 T; 0 U; 4 Other;

Query Match 14.2%; Score 85.4; DB 4; Length 714;
Best Local Similarity 57.8%; Pred. No. 1.3e-14;
Matches 152; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 339 CCGTGTGATGTTCCATCAGTGGATTAACCTTTGACCGGAATGTTATCCAGCTGT 398
DB 1 CGGCGGTGAGCTTCCACGAGTGAGACTTCACTTCAGCGCGGCTGGCCGGCC 60
QY 399 TAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGGCCACTGCGCAAGTCAC 458
DB 61 TGTCTGAGTGGCGCGGACCTGCTGCACACGAGCCGTGGTCCCACTGACCGCAAGTCCC 120
QY 459 ATGACGCGGTTAATAATGCTTTGATCATTTTTCAGATTGGAATTTTGGCTGCTGT 518
DB 121 ACGGCGCATCAACCACTGTTGCGCCACTAGCCGACTGCGACTTCTGGCTGCGCTCT 180
QY 519 ATAAATCCTTTGGGATTTTAAACCCCACTTCAAAAACATGATGATGATCAAAAA 578
DB 181 ACGGCGCGCGCGGACCTTACCGCTCCACCTGCGCAGATTCGCGAGGCGCTGGGCGGGA 240
QY 579 TGTGTGATGACAGAGAATATATGA 601
DB 241 TGTGTGACGAGGCGACGCTCTGA 263

RESULT 41
AAH18715
ID AAH18715 standard; cDNA; 1602 BP.
XX
AC AAH18715;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:18986.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
XX
PR 27-AUG-1999; 99JP-00300253.
XX
PR 11-JAN-2000; 2000JP-00118776.
XX
PR 02-MAY-2000; 2000JP-00183767.
XX
PR 09-JUN-2000; 2000JP-00241899.
XX
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX

DR WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
PS Claim 8; SEQ ID NO 18986; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-RT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification; where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX
SQ Sequence 1602 BP; 334 A; 458 C; 491 G; 319 T; 0 U; 0 Other;

Query Match 14.2%; Score 85.4; DB 4; Length 1602;
Best Local Similarity 57.8%; Pred. No. 2e-14;
Matches 152; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 339 CCGTGTGATGTTCCATCAGTGGATTAACCTTTGACCGGAATGTTATCCAGCTGT 398
DB 1 CGGCGGTGAGCTTCCACGAGTGAGACTTCACTTCAGCGCGGCTGGCCGGCC 60
QY 399 TAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGGCCACTGCGCAAGTCAC 458
DB 61 TGTCTGAGTGGCGCGGACCTGCTGCACACGAGCCGTGGTCCCACTGACCGCAAGTCCC 120
QY 459 ATGACGCGGTTAATAATGCTTTGATCATTTTTCAGATTGGAATTTTGGCTGCTGT 518
DB 121 ACGGCGCATCAACCACTGTTGCGCCACTAGCCGACTGCGACTTCTGGCTGCGCTCT 180
QY 519 ATAAATCCTTTGGGATTTTAAACCCCACTTCAAAAACATGATGATGATCAAAAA 578
DB 181 ACGGCGCGCGGACCTTACCGCTCCACCTGCGCAGATTCGCGAGGCGCTGGGCGGGA 240
QY 579 TGTGTGATGACAGAGAATATATGA 601
DB 241 TGTGTGACGAGGCGACGCTCTGA 263

RESULT 42
AABN50404
ID AABN50404 standard; DNA; 60 BP.
XX
AC AABN50404;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:23152.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.

XX WO200210449-A2.
PN 07-FEB-2002.
XX 20-JUL-2001; 2001WO-1B001903.
XX 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX (COMP-) COMPUGEN INC.
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
XX Example 1; SEQ ID NO 23152; 47bp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcriptome units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridizing selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcripts. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rat, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 60 BP; 22 A; 14 C; 14 G; 10 T; 0 U; 0 Other;
SQ
Query Match 10.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 GAAGAGGCGAGAGATCTCAAGAACCTCATCAAGACAGCTCATAGTGGCCATTCT 258
DB 1 GAAGAGGCGAGAGATCTCAAGAACCTCATCAAGACAGCTCATAGTGGCCATTCT 60
RESULT 43
ACF25360
ID ACF25360 standard; DNA; 7095 BP.
XX ACF25360;
XX
XX 22-SEP-2003 (first entry)
XX
XX Rat microtubule associated protein IB gene.
XX spinal cord; neuropathic pain; central sensitisation pain; pain;
XX analgesic; gene therapy; gene; ss.
XX
XX Rattus norvegicus.
XX OS

PN EP1284298-A2.
XX 19-FEB-2003.
XX 26-JUL-2002; 2002EP-00255229.
XX 27-JUL-2001; 2001GB-00018354.
PR 07-FEB-2002; 2002GB-00002883.
XX (WARN) WARNER LAMBERT CO.
XX Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
XX WPI; 2003-543489/52.
XX P-PADB; ABM04816.
XX
XX Use of an isolated gene sequence in the screening of compounds for
PT diagnosing or treating pain.
XX
XX Claim 1; Page 125-127; 188pp; English.
XX
XX The invention relates to a novel isolated gene sequence that is
CC downregulated in the spinal cord of a mammal in response to mechanically
CC distinct first and second models of neuropathic or central sensitisation
CC pain, useful in the screening of compounds for diagnosing or treating
CC pain. A protein encoded by a gene of the invention has analgesic
CC activity. A polynucleotide of the invention may have a use in gene
CC therapy. The gene sequence is useful for preparing a composition for
CC diagnosing or treating pain. The present sequence represents a gene of
CC the invention
XX
XX Sequence 7095 BP; 2124 A; 1856 C; 1799 G; 1316 T; 0 U; 0 Other;
SQ
Query Match 7.9%; Score 47.4; DB 9; Length 7095;
Best Local Similarity 48.7%; Pred. No. 0.008;
Matches 129; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
QY 53 AATTCGAAACCTGCGCCGTTGAGCGACAAAGAGATCTTGGGTAAATGTGTCGAAA 112
DB 1645 AAACCCGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1704
QY 113 TCCATGCGCACACCTTAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 172
DB 1705 ACTCCCTTAAG 1764
QY 173 GTGACCGAG 232
DB 1765 GAAATCAAAAAG 1824
QY 233 AAGACAGTCATCAAGCTGSCCATTTTATAGGAATATCACTTAATCAAGATGAGCTA 292
DB 1825 AAGAGAGCGCAAG 1884
QY 293 GCATTATGAGAGAAATTTAAGAAGA 317
DB 1885 GAAACCCAAAAGAGAGATTAAGAAGA 1909

RESULT 44
ACH78513
ID ACH78513 standard; DNA; 543 BP.
XX ACH78513;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #11708.
XX Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
XX OS

PN US2003194704-A1.
 XX 16-OCT-2003.
 PD
 XX
 PF 03-APR-2002; 2002US-00029386.
 PR 03-APR-2002; 2002US-00029386.
 XX
 XX (PENN/) PENN S G.
 PA (PENK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX
 PI Penn SG, Rank DR, Hanzel DK;
 XX
 DR MPI, 2004-119264/12.
 XX
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX
 PS Claim 15; SEQ ID NO 11708; 80pp; English.

CC The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 XX
 SQ Sequence 543 BP; 137 A; 128 C; 119 G; 159 T; 0 U; 0 Other;

Query Match 7.8%; Score 47; DB 12; Length 543;
 Best Local Similarity 58.0%; Pred. No. 0.0023;
 Matches 83; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 459 ATGACGGGTTAATAATGCTTTGATCATTGTTTCAGATTGTGAATTTTGGCTGCTTGT 518
 DB 1 ACGGGGGCATCAACACGCTTTTAACCACTTGGCCGATGGAGTTCCTCCACCCCTCT 60
 QY 519 AATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGTGATTAACAAA 578
 DB 61 ATATGCTGATGAGACGTGTAGGCCCAACCTCAAGAGGATTTGTGAAGAAATCAATTAAGT 120

QY 579 TGTGTGATGATGAGAACATATGA 601
 DB 121 TGTGTGATGATGAGAACATCTTTAA 143

RESULT 45
 ABQ76615/c
 ID ABQ76615 standard; cDNA; 15790 BP.
 XX
 AC ABQ76615;
 XX
 DT 21-NOV-2002 (first entry)
 XX
 DE C. albicans BAX-associated cDNA fragment SEQ ID 655.

KW Bax; Bax-resistance; cytosstatic; fungicide; immunosuppressive; virucide;
 KW vasotrophic; vaccine; gene therapy; proliferative disorder; cancer;
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 KW neurodegeneration; cell death; ss.
 XX
 OS Candida albicans.
 XX
 PN WO200264766-A2.
 XX
 PD 22-AUG-2002.
 XX
 PF 21-DEC-2001; 2001WO-BP015398.
 XX
 PR 22-DEC-2000; 2000EP-00870318.
 PR 04-JAN-2001; 2001EP-00870002.
 PR 09-JAN-2001; 2001EP-00870003.

XX (JANC) JANSEN PHARM NV.
 XX
 PI Contreras RH, Eberhardt I, Luyten WHMT, Reekmans RJ;
 XX
 DR MPI; 2002-667002/71.
 XX
 DR P-PSDB; ABG93349.

PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.
 XX
 PS Claim 36; Fig 2; 344pp; English.

CC This invention describes a novel nucleic acid representing a synthetic
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
 CC resistant yeast or fungi, identifying, or obtaining and identifying
 CC Candida spp. sequences that are differentially expressed in a pathway
 CC eventually leading to programmed cell death or identifying inhibitors or
 CC inhibitor sequences of Bax-induced cell death. The products of the
 CC invention have cyostatic, fungicide, immunosuppressive, virucide and
 CC vasotrophic activity and can be used in vaccines or for gene therapy. The
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
 CC antisense molecules and antibodies are useful as medicaments or in
 CC preparing a medicament for treating, preventing and/or alleviating
 CC diseases associated with yeast or fungi or proliferative disorders, such
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds
 CC or polypeptides, or the genetically modified organism are useful for
 CC preparing a medicament for modifying the endogenous flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polynucleotide associated with the Bax gene
 CC described in the disclosure of the invention
 XX
 XX

Sequence 15790 BP; 5181 A; 2570 C; 2199 G; 5840 T; 0 U; 0 Other;

Query Match 7.1%; Score 42.4; DB 6; Length 15790;
 Best Local Similarity 55.4%; Pred. No. 0.36;
 Matches 82; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

ALIGNMENTS

```
RESULT 1
US-09-799-451-537
; Sequence 537, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pc_Fl_genes Version 2.0
; SEQ ID NO 537
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (113)..(664)
US-09-799-451-537

Query Match          30.3%; Score 182; DB 4; Length 1175;
Best Local Similarity 58.7%; Pred. No. 6.1e-47;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGGCCACAGATGCTTTAATTCAAAACTGSCCGTTTCAGGCACAAAGAGA 89
DB 102 CTCAGAGACCCATGAGTCTTCAGCTCAAAAGAGCCTGGCACTGCAGAGAGAGAAGC 161
QY 90 TCTTGGGTAAATGCTGTCCAAATCCATGCGCACCTTAATAGACACAAAGTATG 149
DB 162 TACTAGTAAATGAGGGGTGCGCTCTGTGCTCATCTTCTCAATAGATGAGACAAAGCATG 221
QY 150 AGGTGCTGATGAGCTCTCAAGAGTACACAGGAGATACCCCAAAACAAGAGAGGAG 209
DB 222 AGGTGCTGATGAGCTCTCAAGAGTACACAGGAGATACCCCAAAACAAGAGAGGAG 281
QY 210 AGAAGATCATCAAGAACTCATCAAGAGTACATCAAGTGGCCATCTTTATAGGATA 269
DB 282 AGCGCGTATCAAGAGCTGATCAAAAGTGGCCATCAAGTGGCTGTGCTGCACCGCAAG 341
QY 270 ATCAGTTTATCAAGATGAGCTAGCTTGTATGAGAAATTTAAGAAAGTTATCATCAG 329
DB 342 GCTCCTTGGCCCAAGTGAAGTGGCCCTGCTAACCCGCTTGCAGAAAGTGGCCAGG 401
QY 330 TTGCTATAGACCGGTGCTCAGTTTCCATCAGTGGATTATACCTTTGACCGGAATGTGTAT 389
DB 402 GTGCATAGACGGCATCTTACTTGTGTAGAGTAACTTCACTACCTTCGAGGCTGCTGTCTGG 461
QY 390 CCAGGCTGTAAATGAATGACAGAGATGCTGCACCAAAATCATTCAGCGCCACCTCATG 449
```

```
DB 462 CTGGCCTGTGACCCAGATGCGCGGATGCTGCTAGAGTGGTGAAGACCACTCACGC 521
QY 450 CCAAGTCACATGAGACGGCTTAATATGCTTGTGATCATTTTTCAGATTTGTAATTTGG 509
DB 522 CCAAGTCACATGAGCGGATCCGCAAGTGTGTGATCATCTTCTTACCCAGGTGCTCA 581
QY 510 CTGGCTGTATATCTTTTGGGAATTTTAAACCCCACTTACAAACATATGTAGATGTA 569
DB 582 CGGCCCCCTTATGGGCT-----GACTTCATCTACAGACCTTGGCAAGATCTGTAGCGAC 635
QY 570 TCACAAATATGTTGATGAGAGAGAAATATGA 601
DB 636 TCAGAGAGCTGTACAGCAAGAGGAGCTCTGA 667

RESULT 2
US-09-513-999C-14427
; Sequence 14427, Application US/0951399C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14427
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 139..feature
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 166..feature
; OTHER INFORMATION: r=a or g
; NAME/KEY: misc_feature
; LOCATION: 170
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 208..feature
; OTHER INFORMATION: k=g or t
; NAME/KEY: misc_feature
; LOCATION: 221
; OTHER INFORMATION: r=a or g
US-09-513-999C-14427

Query Match          26.5%; Score 159.2; DB 4; Length 340;
Best Local Similarity 98.8%; Pred. No. 4.7e-40;
Matches 158; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 45 ATGCTTTTATTCAAAAAAGCTGCGCTTCAGGCAAAAGAAATCTTGGTAAATGG 104
DB 181 ATGCTTTTATTCAAAAAAGCTGCGCTTCAGGCAAAAGAAATCTTGGTAAATGG 240
QY 105 TGTCAAAATTCATGCGCCACCTTAATAGACGACAAAGTATGATGAGTGTGATGAGC 164
DB 241 TGTCAAAATTCATGCGCCACCTTAATAGACGACAAAGTATGATGAGTGTGATGAGC 300
QY 165 TCTACAGATGACGAGGAGTACCCCAAAACAAGAGA 204
```


Db 301 TCTACAGAGTGACCCAGGAGTACACCAAAACAAGAGGA 340

```

RESULT 3
US-09-513-999C-14434
: Sequence 14434, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J. B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J. Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
Patent No. 6783961
FILE REFERENCE: 59. US2. REG
CURRENT APPLICATION NUMBER: US/09/513. 999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ. ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 14434
LENGTH: 252
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 32
OTHER INFORMATION: y=c or t
US-09-513-999C-14434

```

Query Match	21.0%	Score 126.4	DB 4	Length 252
Best Local Similarity	98.6%	Pred. No. 9.1e-30		
Matches 138	Conservative 0	Mismatches 1	Indels 1	Gaps 1
QY	1	CTTGAGCGCTCCGGGCGCGCTGC-C-GA	CTCTCCGATGCGCACAAGATGCTTTAATTC	59
Db	113	CTTAGAGGTCCCGGGCGCGCTCGCGC	CACTCTCGATGGCACAATGCTTTAATTC	172
QY	60	AAAACTGCGCGTTTCAGGCACAAAAGA	ATCTTGCTTAAATGTGTCCAATCATCG	119
Db	173	AAACCTGCGCGTTTCAGGCACAAAAGA	ATCTTGCTTAAATGTGTCCAATCATCG	232
QY	120	CCACCACTTATATAGACGAC	139	
Db	233	CCACCACTTATATAGACGAC	252	

RESULT 4
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: pTZgpt-F18
 ;
 ;
 US-08-232-463-14

[illegible]

RESULT 5
 US-09-248-796A-4857
 Sequence 4857, Application US/09248796A
 Patent No. 6747137
 GENERAL INFORMATION:
 APPLICANT: Keith Weinstock et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
 TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.132
 CURRENT APPLICATION NUMBER: US/09/248,796A
 CURRENT FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: US 60/074,725
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 4857


```
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (10398)..(10398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 6.1%; Score 36.8; DB 4; Length 1664976;
Best Local Similarity 48.1%; Pred. No. 37;
Matches 104; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

OY      216  TCATCAGAACCCGTATCAAGCAGTCATCAAGCTGGCCATCTCTTAATAGAAATACAGT 275
Db      1399648  TCTTTAATTACCTCTCATCATTAAGTCAGCTTTTCTTTTGTTTATATAGCTTCAATTCCTT 1399589
OY      276  TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGTTTCATCAGCTTGCTA 335
Db      1399588  ATTTTTCATATVCTTAATCTTATATATATGCAATCTTAGAGAAATTTCAATTAACCTTGTA 1399529
OY      336  TGACCGTGTGATGTTCCATCAGGTGATTTATACCTTTGACCGGAATGTATACCAAGC 395
Db      1399528  TAACTTTTTTACACTCTCCAAATTTTCTTAATATCTAAGTACTTAAGTTCAATTTTGTAGA 1399469
OY      396  TGTAAATGAATGACGAGAGATGCTGCACCAATCA 431
Db      1399468  TTTTAAATATTTCTGCTTATCCAGCACTTGCTCA 1399433

RESULT 9
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Buit et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28252)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature

```

LOCATION: (98266) .. (98266)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (98343) .. (98343)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (103598) .. (103598)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (163385) .. (163385)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (191989) .. (191989)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (231980) .. (231980)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234187) .. (234187)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234220) .. (234220)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234814) .. (234814)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (309398) .. (309398)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (309418) .. (309418)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (312837) .. (312837)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (312993) .. (312993)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (319226) .. (319226)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559167) .. (559167)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559241) .. (559241)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (600992) .. (600992)	a, t, c, or g

OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (622708) .. (622708) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (657081) .. (657081) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (657203) .. (657203) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (674435) .. (674435) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (682442) .. (682442) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (713652) .. (713652) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (741684) .. (741684) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (779455) .. (779455) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (779676) .. (779676) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (855539) .. (855539) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (871619) .. (871619) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (1084630) .. (1084630) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (1096846) .. (1096846) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (1119881) .. (1119881) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (1130881) .. (1130881) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (1310988) .. (1310988) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (1313224) .. (1313224) OTHER INFORMATION: n equals a, t, c, or g	FEATURE: NAME/KEY: misc_feature LOCATION: (1349473) .. (1349473) OTHER INFORMATION: n equals a, t, c, or g
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

	Query Match	Best Local Similarity	4.1%;	Score 36.8;	DB 4;	Length 1664976;
	Matches	104;	Conservative	0;	Mismatches 112;	Indels 0;
					Gaps 0;	
Qy	216	TCATCAAGAACCTTCATCAAGACAGCTCATCAAGCTGGCCATCTTTATATAGAAATATACGT	275			
Db	1399648	TCCTTAATATTCCTCATCAATAAAGTCAGCTTTTCTTTGGTTTATATAGCTTCATTCCTT	1399589			
Qy	276	TTATCAAGATGACCTGACATGATGAGAAATTTAGAGAAAGTCATCAGCTGCTCA	335			
Db	1399588	ATTATTTTCATTTCTTAATCTTATATATGCACTCTTAAGAGAAATTTCAATTAACHTSTTA	1399529			
Qy	336	TGACCGTGGTCAGTTTCCATCAAGCTGATTAATACCTTTGACCGGAATGTGTTATCCAGC	395			
Db	1399588	TAACTTTTTTACACTCTCCAAATTTTCTTAATATCTATGTACTTAAGTTCATTTTTSAGA	1399469			
Qy	396	TGTTAATGATGACGAGAGATGCTGACCAATCA	431			
Db	1399468	TTTTTAATTAATTTTGCTTATTCAGCACTTGCTCA	1399433			

```

RESULT 10
US-09-621-976-2813
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins
FILE REFERENCE: GENSET, 054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ. ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 235..399
US-09-621-976-2813

```

Query Match	6.1%;	Score 36.6;	DB 4;	Length 832;
Best Local Similarity	11.6%;	Pred. No. 0.46;		
Matches	42;	Conservative 165;	Mismatches 154;	Indels 2;
				Gaps 1;

Oy	241	TATCAGGTGGCCATCTCTTATAGAAATACATCACTTAACTAAGATAGCAGATGAT	300
Db	1	YRATWTKTTTAAKCTTKKMSMSYMMWTKYTHKTTTWRNRKKKKAAWKTKTWTWYW	60
Oy	301	GGAGAAATTTAAGAAGAAAGTTCACTGAGCTGTATGACCGTGTCACTTCCATCAGGT	360
Db	61	RYAMWGTGKKKAMCRPTTKKKKKGGYMMWYWGRRSYMAMTTWTGYAAYRRSMYWR	120
Oy	361	GGATTATACCTTTGACCGGAATGKG--TTATCCAGGCGTTAAATGATGACAGAGATG	418
Db	121	YRCHKKAAYRKTTCTYSSKSGTTHMKRKKATTTWTKTYMAATRYMMWMTTKRRPAS	180
Oy	419	CTGCACCAATCATTCAGCGCCACCTCACTGCCAAGTCACATGACGGGTTAATATGTC	478
Db	181	WMYCMWNGKARKSTWKKRSRYASARSAKCYCSGAGMSWKMMWRMWRGATGAGM	240
Oy	479	TTTGATATTTTTCAGATTGGAAATTTTGGCGCCCTTGATATATCTTTTGGAAATTTT	538
Db	241	KAWASCMWRKRYAGSKSTYSKSMWCMWTSWTKCYTKAWWTGYCYRKGSMGKRGRWY	300
Oy	539	AAACCCCACTTACAAAATATGTGATGATGATCAACAAAATGTTGGATGAAGACACATA	598
Db	301	ASKKYTKMKRMMWCARARYSTGTGRASMMWRMYTTMMKMKWYKAMAPAAARMAMWAW	360


```
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7337
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-7337

Query Match
Best Local Similarity 51.6%; Score 36.2; DB 4; Length 502;
Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 386 TTATCCAGCGCTGTTAATGATGAGAGATGCTGACCAATCATTCAGCCGACCTC 445
DB 398 TTGACCATGATTTAATTAATCAATTCATTCGACATTAATACGTTGCGCAACAAAG 339
QY 446 ACTGCCAAGTCACATGAGCGGGTTAATATATGCTTTGATCATTTTTCAGATTGGAATTT 505
DB 338 CCAAGTAAACCAAGATATGTTGTTGTCATTTTGTATTTTGGTTGCTTGTTG 279
QY 506 TTGGCTGCGCTTGTATTAATCCTTTTGGGAATTTAAACCCCA 546
DB 278 GTTGATTAATCTGTAATACGTTTGTGACGCTGAACCGCA 238

RESULT 15
US-09-270-767-22619/c
; Sequence 22619, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22619
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-22619

Query Match
Best Local Similarity 51.6%; Score 36.2; DB 4; Length 502;
Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 386 TTATCCAGCGCTGTTAATGATGAGAGATGCTGACCAATCATTCAGCCGACCTC 445
DB 398 TTGACCATGATTTAATTAATCAATTCATTCGACATTAATACGTTGCGCAACAAAG 339
QY 446 ACTGCCAAGTCACATGAGCGGGTTAATATATGCTTTGATCATTTTTCAGATTGGAATTT 505
DB 338 CCAAGTAAACCAAGATATGTTGTTGTCATTTTGTATTTTGGTTGCTTGTTG 279
QY 506 TTGGCTGCGCTTGTATTAATCCTTTTGGGAATTTAAACCCCA 546
DB 278 GTTGATTAATCTGTAATACGTTTGTGACGCTGAACCGCA 238
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23836
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-23836

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 601;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 184 GTACACCCAAAACAGAGAGGAGAGAGATCATCAAGACCTCATCAAGACAGTCAT 243
DB 202 GAACACACCATTAATAAAGTAGGCAAGATATTAACAGTCACTTCTCAAAAGACAT 143
QY 244 CAAGCTGCCATTCTTATAGAAATATCATGTTAATCAAGATGACATGATGATGGA 303
DB 142 TTATGCGGCAACAAACATATGAGAAAAGCTCAACATCGCTGATCTTAGAGAAATT 83
QY 304 GAAATTTAAGAGAAAGTTCATCAGCTGCTAT 336
DB 82 TATACCTATGACCCCAATTCACCACCATCATAT 50
```

```
RESULT 17
US-09-949-016-173088/c
; Sequence 173088, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173088
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173088

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 601;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 184 GTACACCCAAAACAGAGAGGAGAGATCATCAAGACCTCATCAAGACAGTCAT 243
DB 202 GAACACACCATTAATAAAGTAGGCAAGATATTAACAGTCACTTCTCAAAAGACAT 143
QY 244 CAAGCTGCCATTCTTATAGAAATATCATGTTAATCAAGATGACATGATGATGGA 303
DB 142 TTATGCGGCAACAAACATATGAGAAAAGCTCAACATCGCTGATCTTAGAGAAATT 83
QY 304 GAAATTTAAGAGAAAGTTCATCAGCTGCTAT 336
```

DB 82 TATACCTATGACCCCAATTCACACATCATAT 50

RESULT 18
US-09-540-236-1642/c
; Sequence 1642, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; PRIORITY FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1642
; LENGTH: 630
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-1642

Query Match
Best Local Similarity 60.8%; Score 36.2; DB 4; Length 630;
Pred. No. 0.52;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 281 CAAGATGAGCTAGCTATGAGAAATTAAAGAAAGTTATCATGCTTGTATGACC 340
DB 439 CTACATGTGCAAGCATTTGCTGGACATCTCATTAAGTGGCTGCATGACGACACCTA 380
QY 341 GTGGTCAGTTTCATCAGGTGATTTATACCTTTGACC 377
DB 379 ATAATAATTGGCTCTAGTGATCATGCCCTTGATC 343

RESULT 19
US-09-596-002-15
; Sequence 15, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Paterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; PRIORITY FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 28626
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 15
; PUBLICATION INFORMATION:
US-09-596-002-15

Query Match
Best Local Similarity 60.8%; Score 36.2; DB 4; Length 28626;
Pred. No. 5.1;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 281 CAAGATGAGCTAGCTATGAGAAATTAAAGAAAGTTATCATGCTTGTATGACC 340
DB 17972 CTACATGTGCAAGCATTTGCTGGACATCTCATTAAGTGGCTGCATGACGACACCTA 18031
QY 341 GTGGTCAGTTTCATCAGGTGATTTATACCTTTGACC 377
DB 18032 ATAATAATTGGCTCTAGTGATCATGCCCTTGATC 18068

RESULT 20
US-09-949-016-16647
; Sequence 16647, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16647
; LENGTH: 32039
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16647

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 32039;
Pred. No. 5.4;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 184 GTACACCCCAAAAGAGAGGAGAGAGATCATCAAGACCTCATCAGACAGTCAT 243
DB 17364 GAAACAACACCATCAAAAAGTAGGCAAGATTTAAAGCTGCTCTCTCAAAAAGAGACT 17423
QY 244 CAAGCTGGCCATCTTTATAGAAATATCATGTTATATCAAGATGAGTACATTGATGA 303
DB 17424 TTATGTGGCCCAAAACATATGAGAAAAGCTCAACATCGCTGATCTTTAGAGAAATT 17483
QY 304 GAAATTAGAGAAAGTTCATCAGCTTGCTAT 336
DB 17484 TATACCTATGACCCCAATTCACACCATCATAT 17516

RESULT 21
US-09-949-016-12028
; Sequence 12028, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12028
; LENGTH: 32043
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12028

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 32043;
Pred. No. 5.4;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 184 GTACACCCCAAAAGAGAGGAGAGAGATCATCAAGACCTCATCAGACAGTCAT 243

Db 17368 GACACACACATCAAAAGTAGGCAAGAGATATTAACAGTCATCTTCCAAAAGAGCAT 17427
QY 244 CAAGCTGGCCATCTTTATAGGAATTAATCACTTTAATCAAGATGAGCTAGCATTTGATGA 303
Db 17428 TTATGTGGCCACAAACATATAGAGAAAAGCTCAACATGCTGCTATCTTTAGAGAAAT 17487
QY 304 GAAATTTAAGAGAAAGTTGATCATCGCTTCTAT 336
Db 17488 TATACCTATGACCCCAATTCACCACTCATATAT 17520

RESULT 22

US-09-949-016-17032/c
; Sequence 17032, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17032
; LENGTH: 194937
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17032

Query Match 6.0%; Score 35.8; DB 4; Length 194937;
Best Local Similarity 48.7%; Pred. No. 21;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 71 GTTCAGGCACAAAGAGATCTTGGGTAATGCTGTCCAAATCCATCGCACACCTTA 130
Db 97380 GTGAAGAGACAGCCACAGACTGGAGAGATATTGCCAACTACCCATCTGGCAAGGA 97321
QY 131 ATAGACGACACAAGTAGTGCTGATGATGAGCTTACAGAGTACAGGAGTACACC 190
Db 97320 ATATATACCAAGATATGTATAGAGCTCAACCACTCTGCAGAAAAAACCCTAATATCC 97261
QY 191 CAAAACAGAGAGAGGAGAGAAAGATCATCAAGAACTCATCAAGACATCATCAAGCTG 250
Db 97260 TATTTAAAAAACAATGCAAAAGATGTGAATCGACGTTAAAAAAGACATCAAAATG 97201
QY 251 GCCATCTTTATAGAAATA 269
Db 97200 GCAAAACAGGCATATGAAAA 97182

RESULT 23

US-09-949-016-17033/c
; Sequence 17033, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17033
; LENGTH: 194937
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17033

Query Match 6.0%; Score 35.8; DB 4; Length 194937;
Best Local Similarity 48.7%; Pred. No. 21;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 71 GTTCAGGCACAAAGAGATCTTGGGTAATGCTGTCCAAATCCATCGCACACCTTA 130
Db 97380 GTGAAGAGACAGCCACAGACTGGAGAGATATTGCCAACTACCCATCTGGCAAGGA 97321
QY 131 ATAGACGACACAAGTAGTGCTGATGATGAGCTTACAGAGTACAGGAGTACACC 190
Db 97320 ATATATACCAAGATATGTATAGAGCTCAACCACTCTGCAGAAAAAACCCTAATATCC 97261
QY 191 CAAAACAGAGAGAGGAGAGAAAGATCATCAAGAACTCATCAAGACATCATCAAGCTG 250
Db 97260 TATTTAAAAAACAATGCAAAAGATGTGAATCGACGTTAAAAAAGACATCAAAATG 97201
QY 251 GCCATCTTTATAGAAATA 269
Db 97200 GCAAAACAGGCATATGAAAA 97182

RESULT 24

US-09-248-796A-11532
; Sequence 11532, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 11532
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-11532

Query Match 5.9%; Score 35.4; DB 4; Length 603;
Best Local Similarity 51.6%; Pred. No. 0.9;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 129 TATAGACGACCAAGTAGTGCTGATGAGCTTACAGAGTACAGGAGAGTACA 188
Db 309 TGATATACACAAATAGATGTAAGCAAAAGCATTAAGTGAATATGATTTAAGAGAAAT 368
QY 189 CCCAAACAGAGAGAGGAGAGAGATCATCAAGAACTCATCAAGACAGTCAAGC 248
Db 369 AGCAACAAAGAAAGCTTATAGTCACTTTTGAATATCATAGTCAGATATGTTTGGACT 428
QY 249 TGCCATCTTTATAGGAATATCACTTATATCAAA 285
Db 429 TGGTCATGATTCATATGAAATGTTCACTGGCATTAGA 465

RESULT 25

US-09-167-109-3
; Sequence 3, Application US/09167109

Patent No. 6399297
; GENERAL INFORMATION:
; APPLICANT: Baker, Brenda F.
; APPLICANT: Comsert, Lex M.
; APPLICANT: Monia, Bret P.
; APPLICANT: Xu, Xiaoxing S.
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRAF EXPRESSION
; FILE REFERENCE: ISPH-0321
; CURRENT APPLICATION NUMBER: US/09/167,109
; PRIORITY FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 3
; LENGTH: 2455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)..(1924)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U21092
; DATABASE ENTRY DATE: 1995-03-23
US-09-167-109-3

Query Match
Best Local Similarity 51.6%; Score 35.4; DB 3; Length 2455;
Pred. No. 2.1;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 72 TTCAGGACCAAAAGAGATCTGGTAAATGTGTCAAAATCCATCGCCACACCTTAA 131
DB 1122 TTGAGGACCAAAAGAAATGCTTGAATATGAATCAAAATCTTCAATTACAGCGAG 1181
QY 132 TAGACGACACAAGTAGTGCTGATGAGCTCTACAGAGTACAGGAGTACACCC 191
DB 1182 TGATGACAGCCAGCAGAGAAACTGAAGAGCTTGACAGAGGATCCGGCCCTTCGGC 1241
QY 192 AAAACAGAGAGGAGAGAGATCATCAAGACCT 228
DB 1242 AGAACTGGAGAGCAGACAGCATGAGAGCAGCCT 1278

RESULT 26
US-09-949-016-468
; Sequence 468, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 468
; LENGTH: 2455
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-468

Query Match
Best Local Similarity 51.6%; Score 35.4; DB 4; Length 2455;
Pred. No. 2.1;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 72 TTCAGGACCAAAAGAGATCTGGTAAATGTGTCAAAATCCATCGCCACACCTTAA 131
DB 1122 TTGAGGACCAAAAGAAATGCTTGAATATGAATCAAAATCTTCAATTACAGCGAG 1181

QY 132 TAGACGACACAAGTAGTGCTGATGAGCTCTACAGAGTACAGGAGTACACCC 191
DB 1182 TGATGACAGCCAGCAGAGAACTGAAGAGCTTGACAGAGGATCCGGCCCTTCGGC 1241
QY 192 AAAACAGAGAGGAGAGAGATCATCAAGACCT 228
DB 1242 AGAACTGGAGAGCAGACAGCATGAGAGCAGCCT 1278

RESULT 27
US-08-813-323C-4
; Sequence 4, Application US/08813323C
; Patent No. 6822070
; GENERAL INFORMATION:
; APPLICANT: Baltimore, David
; APPLICANT: Cheng, Genhong
; APPLICANT: Ye, Zheng-sheng
; APPLICANT: Lederman, Seth
; APPLICANT: Cleary, Aileen
; TITLE OF INVENTION: Truncated Craf-1 Inhibits CD40 Signalling
; FILE REFERENCE: 0575/50659
; CURRENT APPLICATION NUMBER: US/08/813,323C
; PRIORITY FILING DATE: 1997-03-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2455
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-08-813-323C-4

Query Match
Best Local Similarity 51.6%; Score 35.4; DB 4; Length 2455;
Pred. No. 2.1;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 72 TTCAGGACCAAAAGAGATCTGGTAAATGTGTCAAAATCCATCGCCACACCTTAA 131
DB 1122 TTGAGGACCAAAAGAAATGCTTGAATATGAATCAAAATCTTCAATTACAGCGAG 1181
QY 132 TAGACGACACAAGTAGTGCTGATGAGCTCTACAGAGTACAGGAGTACACCC 191
DB 1182 TGATGACAGCCAGCAGAGAAACTGAAGAGCTTGACAGAGGATCCGGCCCTTCGGC 1241
QY 192 AAAACAGAGAGGAGAGAGATCATCAAGACCT 228
DB 1242 AGAACTGGAGAGCAGACAGCATGAGAGCAGCCT 1278

RESULT 28
US-09-949-016-4366
; Sequence 4366, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4366
; LENGTH: 7571
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4366

Best Local Similarity 46.8%; Pred. No. 19;
Matches 111; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 80 CAAAAGAGATCTGGGTAATAATGTGTCCAAATCCATGCCACCACTTAATAGAGAC 139
DB 89410 CCAAGAGAAAGAGTGGCTAAAAAGAGCAAAACCTATCAAGAGAGGAAAAACCA 89469
QY 140 ACAAGTAGAGAGTGTGTGATGAGCTCTACAGAGTACAGAGGAGTACACCCAAACAG 199
DB 89470 AAAAGAGAGAGTGAAGAGAGAGTCAAAAAGAGATCAAGAGAGAGAGAGAGAGAA 89529
QY 200 AAGAGGACAGAGAGATCATCAAGAACTCATCAAGACATCATCAAGCTGGCATTCTT 259
DB 89530 CCCAAGAAAGAGGTTAAGAGAGAAACCGCCAAAGAGAGTCAAGAGAGAGTTAAGAG 89589
QY 260 TATAGAAATATCATGTTTATATCAAGATGAGTACATTTAGAGAAATTTAAGAG 316
DB 89590 GAAAGAAAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 89646
RESULT 32
US-09-949-016-17590
; Sequence 17590, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17590
; LENGTH: 247299
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17590
Query Match 5.9%; Score 35.4; DB 4; Length 247299;
Best Local Similarity 56.4%; Pred. No. 33;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 163 GCTCTACAGAGTACAGGAGGATACCCCAAAACAGAGAGGACAGAGATCATCA 222
DB 180916 GCTCTACATTTACTTAACAAAGAACTCGAAATTTGTGATGAATTAATAAAGAA 180975
QY 223 GAACCTCATCAAGACATCATCAAGCTGGCATTCTTTATAGAAATATCATGTTAA 279
DB 180976 GAACCTCATCATATCATTCATCATCTGCCATTGTTTATTTGAGTGTGCTTAA 181032
RESULT 33
US-09-902-540-8128
; Sequence 8128, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8128
; LENGTH: 2455
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8128
Query Match 5.9%; Score 35.2; DB 4; Length 2455;
Best Local Similarity 49.5%; Pred. No. 2.4;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 149 GAGTGTGATGATGAGCTCTACAGAGTGAACAGGAGTACACCCAAACAGAGAGGCA 208
DB 616 GAGAGGCTCTACAGAGCTGATGAGAGGTGATGAGATTCTCAGGTGAGAGAGATC 675
QY 209 GAGAGATCATCAAGACCTCATCAAGACATCATCAAGCTGGCATTCTTTATAGAA 268
DB 676 CCACGCGCGTCAAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 735
QY 269 AATCAGTTTATCAAGATGAGCTACATTTGATGAGAGAAATTTAAGAGAAAGTTCATCAG 328
DB 736 ATGCAAGGCTTCAAGAGAGAGCTGGGTGAGCGCGAGAGTTCAGAGAGAGATTCAGAG 795
QY 329 CTTG 332
DB 796 ATTG 799
RESULT 34
US-09-902-540-832/c
; Sequence 832, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 832
; LENGTH: 8352
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-832
Query Match 5.9%; Score 35.2; DB 4; Length 8352;
Best Local Similarity 49.5%; Pred. No. 3;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 149 GAGTGTGATGATGAGCTCTACAGAGTGAACAGGAGTACACCCAAACAGAGAGGCA 208
DB 4610 GAGAGCTCTACAGAGCTGATGAGAGGTGAGATTCAGATTCTCAGGTGAGAGAGATC 4551
QY 209 GAGAGATCATCAAGACCTCATCAAGACATCATCAAGCTGGCATTCTTTATAGAA 268
DB 4550 CCACGCGCGTCAAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4491
QY 269 AATCAGTTTATCAAGATGAGCTACATTTGATGAGAGAAATTTAAGAGAAAGTTCATCAG 328
DB 4490 ATGCAAGGCTTCAAGAGAGAGCTGGGTGAGCGCGAGAGTTCAGAGAGAGATTCAGAG 4431
QY 329 CTTG 332
DB 4430 ATTG 4427
RESULT 35
US-09-949-016-13146

```

; Sequence 13146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13146
; LENGTH: 58829
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(58829)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13146

```

```

Query Match
Best Local Similarity 5.8%; Score 35.2; DB 4; Length 58829;
Matches 67; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```

```

QY 467 GTTAAATAGTCTTGTGATCATTTTTCAGATTGGAATTTTGGCTGCTGTATATCCT 526
DB 57676 GTTGTGATATCTCTATGACATTTTAACTTTAGATTGTGAATGACATGTGTAAT 57735
DB 57736 TTTGTTTATTAACCAAGATTATTAACAAGTTTAAATGTGTCAACAACCTGTTCCAT 57795

```

```

RESULT 36
US-09-949-016-194263
; Sequence 194263, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194263
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-194263

```

```

Query Match
Best Local Similarity 5.8%; Score 35; DB 4; Length 601;
Matches 95; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 93 TGGGTAATGTGTCCTCAATTCATCCGACCACTTATAGACGACACAAGTAGAGG 152
DB 231 TGTCTATATTATATCTAATCATGAGGCTCCGATATATTAAGCAATATGAGAGAG 230

```

```

QY 153 TGCTGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGAGGACAGAGA 212
DB 291 TGAAGCAAGATGATACAGACAGATATATATATCAACTTAAGACCCCATTTGCAATAC 350
QY 213 AGATCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGCGCATTTCTTATAGAAATATC 272
DB 351 AGATAGAAAATATACACAGAGATCATATTAAGAAAGCTTGGGCAACATTACAGCAATC 410
QY 273 AGTTAATCAAGATG 287
DB 411 AATTAAACAGACATG 425

```

```

RESULT 37
US-09-949-016-17262/c
; Sequence 17262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17262
; LENGTH: 24428
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17262

```

```

Query Match
Best Local Similarity 5.8%; Score 35; DB 4; Length 24428;
Matches 95; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

```

QY 93 TGGCTAAATGTGTCCTCAATTCATCCGACCACTTATAGACGACACAAGTAGAGG 152
DB 15975 TGTCTATATTATATATATTAATTAATCAATGAGGCTCCGATATATTAAGCAATATGAGAGAG 15916
QY 153 TGCTGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGAGGACAGAGA 212
DB 15915 TGAAGCAAGATGATACAGACAGATATATATTAATCAACTTAAGACCCCATTTGCAATAC 15856
QY 213 AGATCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGCGCATTTCTTATAGAAATATC 272
DB 15855 AGATAGAAAATATACACAGAGATCATATTAAGAAAGCTTGGGCAACATTACAGCAATC 15796
QY 273 AGTTAATCAAGATG 287
DB 15795 AATTAAACAGACATG 15781

```

```

RESULT 38
US-09-949-016-14033/c
; Sequence 14033, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14033
; LENGTH: 784019
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(784019)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14033
```

```

Query Match
Best Local Similarity 5.8%; Score 35; DB 4; Length 784019;
Matches 56; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```

QY 305 AATTAAAGAAAGTTGATCATGCTTGATGACCGGTGATGTTCCATCAGGTGAT 364
DB 891 AATTAAAGAAAGTTGATCATGCTTGATGACCGGTGATGTTCCATCAGGTGAT 364
QY 365 TATACCTTTGACCGGATGTTGTTATCAGGC 395
DB 831 TATACCTTTAAAGAAAGTTGATGATGATGATGATGATGATGATGATGATGAT 801
```

```

RESULT 39
US-09-949-016-205410/C
; Sequence 205410, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 205410
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-205410
```

```

Query Match
Best Local Similarity 5.8%; Score 34.8; DB 4; Length 601;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```

QY 221 AAGAACTCATCAAGACATGATCAAGCTGGCCATCTTTATAGAAATATCATTTAAT 280
DB 365 AAGGACACCATCAAGAAAGTGAAGAGAGCGCCCAAGTTGAGAGAAATATCTTCAAAAT 306
QY 281 CAAGATGAGCTGATGATGAGAAATTTAAGAAAGAAATT 322
DB 305 CATGATAGCACTTTTGTATGTTAACTTGAGGAAACATTTT 264
```

```

RESULT 40
US-09-949-016-205411/C
; Sequence 205411, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 205411
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-205411
```

```

Query Match
Best Local Similarity 5.8%; Score 34.8; DB 4; Length 601;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```

QY 221 AAGAACTCATCAAGACATGATCAAGCTGGCCATCTTTATAGAAATATCATTTAAT 280
DB 534 AAGGACACCATCAAGAAAGTGAAGAGAGCGCCCAAGTTGAGAGAAATATCTTCAAAAT 475
QY 281 CAAGATGAGCTGATGATGAGAAATTTAAGAAAGAAATT 322
DB 474 CATGATAGCACTTTTGTATGTTAACTTGAGGAAACATTTT 433
```

```

RESULT 41
US-09-949-016-17549
; Sequence 17549, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17549
; LENGTH: 24345
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17549
```

```

Query Match
Best Local Similarity 5.8%; Score 34.8; DB 4; Length 24345;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```

QY 221 AAGAACTCATCAAGACATGATCAAGCTGGCCATCTTTATAGAAATATCATTTAAT 280
DB 13643 AAGGACACCATCAAGAAAGTGAAGAGAGCGCCCAAGTTGAGAGAAATATCTTCAAAAT 13702
QY 281 CAAGATGAGCTGATGATGAGAAATTTAAGAAAGAAATT 322
DB 13703 CATGATAGCACTTTTGTATGTTAACTTGAGGAAACATTTT 13744
```

```

RESULT 42
US-09-908-988B-5
; Sequence 5, Application US/09908988B
; Patent No. 6740751
```


Qy 218 ATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTG 257
| | | | | | | | | | | | | | | | | | | | | |
Db 616 ATGCACAACTCTCTGAGAGACTGACATCGACAGGTGAGTC 655

Search completed: July 28, 2005, 19:03:54
Job time : 187 secs


```

c 81 36.4 6.1 8692 19 US-10-623-500-13 Sequence 13, Appl
c 82 36.4 6.1 98642 13 US-10-087-192-1105 Sequence 1105, Ap
c 83 36.2 6.0 759 20 US-10-425-115-137599 Sequence 137599,
c 84 36.2 6.0 2739 18 US-10-335-977-3199 Sequence 3199, Ap
c 85 36.2 6.0 3135 18 US-10-335-977-3200 Sequence 1500, Ap
c 86 36.2 6.0 28626 18 US-10-672-787-15 Sequence 15, Appl
c 87 36.2 6.0 33206 9 US-09-880-107-2380 Sequence 2380, Ap
c 88 36.2 6.0 40045 21 US-10-741-600-17795 Sequence 17795, A
c 89 36 6.0 828 18 US-10-424-559-109224 Sequence 109224,
c 90 35.8 6.0 272 9 US-09-864-761-18684 Sequence 18684, A
c 91 35.8 6.0 482 9 US-09-864-761-1935 Sequence 1935, Ap
c 92 35.8 6.0 49736 20 US-10-719-993-7050 Sequence 7050, Ap
c 93 35.8 6.0 153170 20 US-10-723-860-199 Sequence 199, Ap
c 94 35.8 6.0 185458 20 US-10-719-993-6864 Sequence 6864, Ap
c 95 35.8 6.0 253661 19 US-10-741-601-5611 Sequence 5611, Ap
c 96 35.8 6.0 261817 13 US-10-087-192-2002 Sequence 2002, Ap
c 97 35.8 6.0 310692 17 US-10-428-487-11 Sequence 11, Appl
c 98 35.8 6.0 310268 19 US-10-367-094-195 Sequence 195, Ap
c 99 35.6 5.9 1174 17 US-10-282-122A-23404 Sequence 23404, A
c 100 35.4 5.9 2455 8 US-08-813-323A-4 Sequence 4, Appl

```

ALIGNMENTS

RESULT 1

```

US-10-627-571-1
; Sequence 1, Application US/10627571
; Publication No. US20040082771A1
; GENERAL INFORMATION:
; APPLICANT: KASID, Usha N.
; APPLICANT: KUMAR, Deepak
; APPLICANT: GOKHALE, Prafulla
; APPLICANT: AHMAD, Imran
; TITLE OF INVENTION: ANTI-APOPTOTIC GENE SCC-S2 AND DIAGNOSTIC AND THERAPEUTIC USES
; FILE REFERENCE: 223316
; CURRENT APPLICATION NUMBER: US/10/627,571
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: US 60/264,062
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US02/02212
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: SCC-S2
US-10-627-571-1

```

```

Query Match          100.0%; Score 601; DB 18; Length 1915;
Best Local Similarity 100.0%; Pred. No. 3.9e-169;
Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CTTGACGCTCCGGCGCGCGCTCCGACCTCTCCGATGGCCACAGATGTCTTTAATTCAA 60
DB 100 CTTGACGCTCCGGCGCGCGCTCCGACCTCTCCGATGGCCACAGATGTCTTTAATTCAA 159
QY 61 AAACCTGGCGCTTCAGGCAACAAAGAGATCTTGGGTAATAATGGTCCAAATCCATCCG 120
DB 160 AAACCTGGCGCTTCAGGCAACAAAGAGATCTTGGGTAATAATGGTCCAAATCCATCCG 219
QY 121 CACCACTTAATAGACGACAAGTAGTAGTGTGATGAGCTCTACAGAGTAGCCAG 180
DB 220 CACCACTTAATAGACGACAAGTAGTAGTGTGATGAGCTCTACAGAGTAGCCAG 279
QY 181 GGAATACACCCAAACAAAGAGAGGCAAGAAATCATCAAGAACCTCATCAAGACAGT 240
DB 280 GGAATACACCCAAACAAAGAGAGGCAAGAAATCATCAAGAACCTCATCAAGACAGT 339

```

```

QY 241 CATCAAGCTGGCCATTCTTTATAGGAATAATCAGTTAATCAAGATGAGTAGCATGTAT 300
DB 340 CATCAAGCTGGCCATTCTTTATAGGAATAATCAGTTAATCAAGATGAGTAGCATGTAT 399
QY 301 GGAGAAATTTAAGAAAGAGATTCATCAGCTTGCTATGACCGGTGTCAGTTCCATCAGT 360
DB 400 GGAGAAATTTAAGAAAGAGATTCATCAGCTTGCTATGACCGGTGTCAGTTCCATCAGT 459
QY 361 GGATTAATCCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGAATGACAGAGATGCT 420
DB 460 GGATTAATCCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGAATGACAGAGATGCT 519
QY 421 GGACCAATTCATTCAGCGCCACCTCAGCCACAGTCACATGACAGGCTTAAATAGTCTT 480
DB 520 GGACCAATTCATTCAGCGCCACCTCAGCCACAGTCACATGACAGGCTTAAATAGTCTT 579
QY 481 TGATCATTTTTCAGATTGGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTTAA 540
DB 580 TGATCATTTTTCAGATTGGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTTAA 639
QY 541 ACCCCACTTACAAAACCTATGTGATGTATCAACAAAATGTTGATGAAGACATATG 600
DB 640 ACCCCACTTACAAAACCTATGTGATGTATCAACAAAATGTTGATGAAGACATATG 699
QY 601 A 601
DB 700 A 700

```

RESULT 2

```

US-10-491-545A-48
; Sequence 48, Application US/10491545A
; Publication No. US20050130117A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Peter
; APPLICANT: Li, Congfen
; APPLICANT: Liao, X. Charlene
; APPLICANT: Masuda, Esteban
; APPLICANT: Pardo, Jorge
; APPLICANT: Zhao, Haoran
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modulators of Lymphocyte Activation and Migration
; FILE REFERENCE: 021044-000330US
; CURRENT APPLICATION NUMBER: US/10/491,545A
; CURRENT FILING DATE: 2004-04-01
; PRIOR APPLICATION NUMBER: US 60/327,212
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: WO PCT/US02/31618
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human TNF-induced protein G2-1 mRNA,
; OTHER INFORMATION: complete CDS
US-10-491-545A-48

```

```

Query Match          97.9%; Score 508.4; DB 22; Length 1892;
Best Local Similarity 99.7%; Pred. No. 2.3e-165;
Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 1 CTTGACGCTCCGGCGCGCGCTCCG-CGACTCCCTCCGATGGCCACAGATGTCTTTAATTC 59
DB 63 CTTGACGCTCCGGCGCGCGCTCCG-CGACTCCCTCCGATGGCCACAGATGTCTTTAATTC 122
QY 60 AAACCTGGCGCTTCAGGCAACAAAGAGATCTTGGGTAATAATGGTGTCCAAATCCATCG 119
DB 123 AAACCTGGCGCTTCAGGCAACAAAGAGATCTTGGGTAATAATGGTGTCCAAATCCATCG 182

```

```

; PRIOR APPLICATION NUMBER: 60/068,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-097-065-98

Query Match      97.9%; Score 588.2; DB 14; Length 1943;
Best Local Similarity 99.3%; Pred. No. 2.7e-165;
Matches 590; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY      8 GTCCCGGCGCGGTGCGCGCACTCTCCGATGCGCACAGATGCTTTAATCCAAAAAACCCTG 67
DB      93 GTCCCGGCGCGGTGCGCGCACTCTCCGATGCGCACAGATGCTTTAATCCAAAAAACCCTG 152

QY      68 GCCGTTCAAGCACAAGAAAGATCTTTGGTAAATGTGTCCAAATCCATCGCCACAC 127
DB      153 GCCGTTCAAGCACAAGAAAGATCTTTGGTAAATGTGTCCAAATCCATCGCCACAC 212

QY      128 TTATATAGCACAACAAGTAGAGGTGCTGATGATGAGCTTACAGAGTACCAAGGAGTAC 187
DB      213 TTATATAGCACAACAAGTAGAGGTGCTGATGAGCTTACAGAGTACCAAGGAGTAC 272

QY      188 ACCCAAAACAAGAGAAGGAGGAGAGAGATCATCAAGAACTCATCAAGACATCATCAAG 247
DB      273 ACCCAAAACAAGAGAAGGAGGAGAGAGATCATCAAGAACTCATCAAGACATCATCAAG 332

QY      248 CTGGCCATCTTTTATAGAAATTAATCAGTTTATCAAGATGAGCTAGCATTTGATGAGAA 307
DB      333 CTGGCCATCTTTTATAGAAATTAATCAGTTTATCAAGATGAGCTAGCATTTGATGAGAA 392

QY      308 TTTAAGAAAGAAAGTTTCATCAGCTTGCTATGACCGTGTCAAGTTTTCATCAGGTGATTAT 367
DB      393 TTTAAGAAAGAAAGTTTCATCAGCTTGCTATGACCGTGTCAAGTTTTCATCAGGTGATTAT 452

QY      368 ACCTTTACCGGAATGTGTTATCCAGGCTGTTAATGAAATCCAGAGAGATGCTGACCAA 427
DB      453 ACCTTTACCGGAATGTGTTATCCAGGCTGTTAATGAAATCCAGAGAGATGCTGACCAA 512

QY      428 ATCATTTAGCGGCACCTCATCTGCCAAGTCATGGAACGGGTAAATATGTCTTTGATCAT 487
DB      513 ATCATTTAGCGGCACCTCATCTGCCAAGTCATGGAACGGGTAAATATGTCTTTGATCAT 572

QY      488 TTTTCAGATTGTGAATTTTGGCTGCTGTTATTAATCTTTTGGAAATTTTAAACCCAC 547
DB      573 TTTTCAGATTGTGAATTTTGGCTGCTGTTATTAATCTTTTGGAAATTTTAAACCCAC 632

QY      548 TTACAAAAAATATGTGATGATTAACAAGAAATGTGATGGAAGAAATATATGA 601
DB      633 TTACAAAAAATATGTGATGATTAACAAGAAATGTGATGGAAGAAATATATGA 686

RESULT 4
US-10-372-876-98
; Sequence 98, Application US/10372876
; Publication No. US20030204071A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/372,876
; CURRENT FILING DATE: 2003-02-26
; PRIOR APPLICATION NUMBER: 09/334,595
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-372-876-98

Query Match          97.9%; Score 588.2; DB 17; Length 1943;
Best Local Similarity 99.3%; Pred. No. 2.7e-165;
Matches 550; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      8 GTCGCCGCCCGTCGCCGACTCTCCGATGCGCACAGATGCTTTATTCAAAAACCTG 67
DB      93 GTCGCCGCCCGTCGCCGACTCTCCGATGCGCACAGATGCTTTATTCAAAAACCTG 152
QY      68 GCCGTTCAAGCACAAAAAGATCTTGGGTAATGCTGTCCAATTCATCGCCACACC 127
DB      153 GCCGTTCAAGCACAAAAAGATCTTGGGTAATGCTGTCCAATTCATCGCCACACC 212
QY      128 TTAATGAGAGACACAAGTAGTGAGTGATGAGCTCTACAGATGAGCCAGGGAGTAC 187
DB      213 TTAATGAGAGACACAAGTAGTGAGTGAGTGAGCTCTACAGATGAGCCAGGGAGTAC 272
QY      188 ACCCAAAAACAAGAGAGGACAGAGATCATCAAGAACTCATCAAGACATCATAG 247
DB      273 ACCCAAAAACAAGAGAGGACAGAGATCATCAAGAACTCATCAAGACATCATAG 332
QY      248 CTGGCCATCTTTATTAAGAAATTAATCAAGTTAATCAAGTAGAGTAGATGAGAAA 307
DB      333 CTGGCCATCTTTATTAAGAAATTAATCAAGTTAATCAAGTAGAGTAGATGAGAAA 392
```

```

QY      308 TTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTGACAGTTTCATCAGGTGATTAT 367
DB      393 TTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTGACAGTTTCATCAGGTGATTAT 452
QY      368 ACCTTGACCGGAATGTGTTATCCAGGCTTTAAATGAATGACAGAGATGTCGACCA 427
DB      453 ACCTTGACCGGAATGTGTTATCCAGGCTTTAAATGAATGACAGAGATGTCGACCA 512
QY      428 ATCATTTACGGCCACCTCTGCTCCAAATGACATGAGACGGGTTAAATATGCTTTGATCAT 487
DB      513 ATCATTTACGGCCACCTCTGCTCCAAATGACATGAGACGGGTTAAATATGCTTTGATCAT 572
QY      488 TTTTCAGATTGGAATTTTTGGCTGCTGTATTAATCTTTGGGAATTTTAAACCCAC 547
DB      573 TTTTCAGATTGGAATTTTTGGCTGCTGTATTAATCTTTGGGAATTTTAAACCCAC 632
QY      548 TTACAAAAAATATGTATGATGATCAACAAATGTTGATGAAGAGAACATATGA 601
DB      633 TTACAAAAAATATGTATGATGATCAACAAATGTTGATGAAGAGAACATATGA 686
```

```

RESULT 5
; Sequence 215, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 215
; LENGTH: 1814
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-755-889-215
```

```

Query Match          93.9%; Score 564.4; DB 19; Length 1814;
Best Local Similarity 99.8%; Pred. No. 3.6e-158;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      36 TGCCACAGATGCTTTAATTCAAAAACCTGGCGTTCCAGGCACAAAAAGAGATCTTGG 95
DB      175 TGCCACAGATGCTTTAATTCAAAAACCTGGCGTTCCAGGCACAAAAAGAGATCTTGG 234
QY      96 GTAATATGCTGTCAAATTCATCGCCACCTTAATAGACGACACAAGTAGTAGTGTC 155
DB      235 GTAATATGCTGTCAAATTCATCGCCACCTTAATAGACGACACAAGTAGTAGTGTC 294
QY      156 TGATGAGCTCTACAGAGTGACCAAGAGATACCCAAAAACAAGAGGCGACAGAGA 215
DB      295 TGATGAGCTCTACAGAGTGACCAAGAGATACCCAAAAACAAGAGGCGACAGAGA 354
QY      216 TCATCAAGAACTCTATCAAGACAGTCATCAAGCTGGCCATTCTTTATGGAATTAATCAGT 275
DB      355 TCATCAAGAACTCTATCAAGACAGTCATCAAGCTGGCCATTCTTTATGGAATTAATCAGT 414
QY      276 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGCTA 335
DB      415 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGCTA 474
QY      336 TGACCGTGCTCAGTTTCATCAGGTGATTAATCTTTGAACCGGAATGTTATCCAGGC 395
DB      475 TGACCGTGCTCAGTTTCATCAGGTGATTAATCTTTGAACCGGAATGTTATCCAGGC 534
```

OY	366	TGTTAAATGAATGCAGAGATGCTGCAACCAATCAATTGAGCCCACTCACTGCAAGT	455
Db	535	TGTTAAATGAATGCAGAGATGCTGCAACCAATCAATTGAGCCCACTCACTGCAAGT	594
OY	456	CACATGACGGGTTAATTAATGTCCTTGATCATTTTCAGATTGTGAATTTTGGCTGCT	515
Db	595	CACATGACGGGTTAATTAATGTCCTTGATCATTTTCAGATTGTGAATTTTGGCTGCT	654
OY	516	TGTTAATATCCTTTGGGAAATTTTAAACCCCACTTACAAAACATGTGATGATCAACA	575
Db	655	TGTTAATATCCTTTGGGAAATTTTAAACCCCACTTACAAAACATGTGATGATCAACA	714
OY	576	AAATGTGATGAAGAACAATATGA	601
Db	715	AAATGTGATGAAGAACAATATGA	740

```

RESULT 6
US-10-755-889-217
; Sequence 217, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755, 889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440, 068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469, 757
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 217
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-217

```

Query Match	93.9%	Score 564.4	DB 19	Length 2003
Best Local Similarity	99.8%	Pred. No. 3.9e-158		
Matches 565	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	36	TGGCCACAGATGTCCTTTAATTCAAAAACCTGGCCGTTCAAGCACAAAAGAGATCTTGG	95	
Db	205	TGGCCACAGATGTCCTTTAATTCAAAAACCTGGCCGTTCAAGCACAAAAGAGATCTTGG	264	
QY	96	GTAATAATGTCCTCCAAATCCATTCGCCACCACTTATATGACGACAAAGTGTGAGTGC	155	
Db	265	GTAATAATGTCCTCCAAATCCATTCGCCACCACTTATATGACGACAAAGTGTGAGTGC	324	
QY	156	TGATAGAGCTCTACAGAGTGACCAAGGAGTCAACCCAAAACAAGAGGACAGAAAGA	215	
Db	325	TGATAGAGCTCTACAGAGTGACCAAGGAGTCAACCCAAAACAAGAGGACAGAAAGA	384	
QY	216	TCATCAGAACCCTCATCAAGACAGTCATCAAGCTGGCCATTCCTTATAGAAATACGT	275	
Db	385	TCATCAGAACCCTCATCAAGACAGTCATCAAGCTGGCCATTCCTTATAGAAATACGT	444	
QY	276	TTAATCAAGATGAGGTAGCATTTGATGGAATAATTAAACAAGATTCATCAGCTTGCTA	335	
Db	445	TTAATCAAGATGAGGTAGCATTTGATGGAATAATTAAACAAGATTCATCAGCTTGCTA	504	
QY	336	TGACCGTGTCAATTCATCAGGTGATATACCTTTGACCGAATGTGTTATCCAGGC	395	
Db	505	TGACCGTGTCAATTCATCAGGTGATATACCTTTGACCGAATGTGTTATCCAGGC	564	
QY	396	TGTTAAATGATGCGAGAGATGCTGCACCAATATATTCACGCGCACTCACTGCAAGT	455	
Db	565	TGTTAAATGATGCGAGAGATGCTGCACCAATATATTCACGCGCACTCACTGCAAGT	624	
QY	456	CACATGACGGGTTAATATGTCCTTTGATCAATTTTCAGATTGTGAATTTTGGCTGCGCT	515	

DB	Sequence	Position
Db	625 CACATGGACCGGTTAAATGATGTTGATCATTTTTCAGATTGGAAATTTTGGCTGCT	68
Oy	516 TGTATATCCTTTTGGGAATTTAAACCCCACTTACAAAACTATGTGATGATCAACA	575
Db	685 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA	744
Oy	576 AATGTGTGATGAAGGAACATATGA 601	
Db	745 AAAATGTTGATGAAGGAACATATGA 770	

```

RESULT 7
US-10-087-192-497
; Sequence 497, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 497
; LENGTH: 2034
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-497

```

Query Match	Similarity	93.9%	Score 564.4	DB 13	Length 2034
Best Local	Similarity	99.8%	Pred. No. 3.9e-158		
Matches 565	Conservative	0	Mismatches 1	Indels 0	Gaps 0
Qy	36	TGGCCACAGATGTCCTTTAAATTC	CAAAAAACCTGGCCGTTTCAGGCA	CAAAAAGAAAGATCTTGG	95
Db	249	TGGCCACAGATGTCCTTTAAATTC	CAAAAAACCTGGCCGTTTCAGGCA	CAAAAAGAAAGATCTTGG	308
Qy	96	GTAATAATGCTGTCCAAATCCATG	CCACCACTTTAATAGACACAAAGTAGT	AGGTGC	155
Db	309	GTAATAATGCTGTCCAAATCCATG	CCACCACTTTAATAGACACAAAGTAGT	AGGTGC	368
Qy	156	TGATGAGCTCTACAGAGTGCAC	AGGAGTACACCCAAAAACAGAGAGG	CAGAGAA	215
Db	369	TGATGAGCTCTACAGAGTGCAC	AGGAGTACACCCAAAAACAGAGAGG	CAGAGAA	428
Qy	216	TCATCAGAAACCTCATTAAGACAT	CTCAAGCTGGCCATTTCTTATAGAA	TATACGT	275
Db	429	TCATCAGAAACCTCATTAAGACAT	CTCAAGCTGGCCATTTCTTATAGAA	TATACGT	488
Qy	276	TTAATCAAGATGAGCTGACATG	TGAGAAATTTAAGAAAGATT	CATCAGCTGCTA	335
Db	489	TTAATCAAGATGAGCTGACATG	TGAGAAATTTAAGAAAGATT	CATCAGCTGCTA	548
Qy	336	TGACCGTGTCAGTTTCCATCAG	GTGATTATACCTTTGACCGGAATGT	TTATCCAGGC	395
Db	549	TGACCGTGTCAGTTTCCATCAG	GTGATTATACCTTTGACCGGAATGT	TTATCCAGGC	608
Qy	396	TGTTAAATGATGACAGAGATG	CTGACCAAAATCATTCAGCGC	CATCTGCCAAGT	455
Db	609	TGTTAAATGATGACAGAGATG	CTGACCAAAATCATTCAGCGC	CATCTGCCAAGT	668
Qy	456	CACATGAGCGGGTTAAATATG	TCTTTGATCATTTTCAGATGGAAT	TTTTTGGCTGCT	515
Db	669	CACATGAGCGGGTTAAATATG	TCTTTGATCATTTTCAGATGGAAT	TTTTTGGCTGCT	728
Qy	516	TGTATATACCTTTTGGGAATTT	TAAACCCACCTTACAAAAACTAT	TGATGATGATCA	575

Db	729	TGTTAAATCCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTATGATGATCAACA	788
Oy	576	AAATGTTGATGAAGAAACATATGA	601
Db	789	AAATGTTGATGAAGAAACATATGA	814

```

RESULT 8
US-10-755-889-39
; Sequence 39, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 2081
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-39

```

Query Match	93.9%	Score 564.4;	DB 19;	Length 2081;
Best Local Similarity	99.8%;	Pred. No. 4e-158;		
Matches 565;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

QY	36	TGGCCACAGATGTCCTTTATTCGAAAACCTGGCCGCTTCAGGACCAAAAAGAAATCTTGG	95
Db	268	TGGCCACAGATGTCCTTTATTCGAAAACCTGGCCGCTTCAGGACCAAAAAGAAATCTTGG	327
QY	96	GTTAAATGCTGTCGCAATCCATCGCCACCACTTAAATGACGACAGAATGAGAGTGC	155
Db	328	GTTAAATGCTGTCGCAATCCATCGCCACCACTTAAATGACGACAGAATGAGAGTGC	387
QY	156	TGGATGAGCTCTACAGAGTGACCCAGGAGTACACCCNAAAAGAGAGGACGAGAGA	215
Db	388	TGGATGAGCTCTACAGAGTGACCCAGGAGTACACCCNAAAAGAGAGGACGAGAGA	447
QY	216	TCATCAAGAACCTTCATCAAGACAGTCATCAAGCTGGCCATCTTTATGAGAAATATCAGT	275
Db	448	TCATCAAGAACCTTCATCAAGACAGTCATCAAGCTGGCCATCTTTATGAGAAATATCAGT	507
QY	276	TTAATCAAGATAGCTAGCATTGATGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGCTA	335
Db	508	TTAATCAAGATAGCTAGCATTGATGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGCTA	567
QY	336	TGACCGTGTCACTTCCATCAAGTGGATTAATCCTTTGACCGGAATGCTTATCCAGGC	395
Db	568	TGACCGTGTCACTTCCATCAAGTGGATTAATCCTTTGACCGGAATGCTTATCCAGGC	627
QY	396	TGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCCGCCACCTCACTGCAGAT	455
Db	628	TGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCCGCCACCTCACTGCAGAT	687
QY	456	CACATGAGACGGGTTAATATGCTTGTGATCATTTTTCAGATGTGTGAATTTTGGCTGCCT	515
Db	688	CACATGAGACGGGTTAATATGCTTGTGATCATTTTTCAGATGTGTGAATTTTGGCTGCCT	747
QY	516	TGTATATATCCTTTTGGGAATTTTAAACCCACCTACAAAAAACATATGATGATGATCAACA	575
Db	748	TGTATATATCCTTTTGGGAATTTTAAACCCACCTACAAAAAACATATGATGATGATCAACA	807
QY	576	AAATGTTGATGAAGAAACATATGA	601

```

DB              808  AAATGTTGATGAAGAGAACATATGA 833

RESULT 9
US-10-087-192-496
; Sequence 496, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ. ID NOS: 2059
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 496
; LENGTH: 58723
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-496

```

Query Match	93.9%	Score 564.4	DB 13	Length 58723
Best Local Similarity	99.8%	Pred. No. 3e-157		
Matches 565; Conservative	0	Mismatches 1	Indels 0	Gaps 0

OY	36	TGGCCACAGATGCTTTAAATCCAAAACCTGGCCGCTTGAGGCACAAAAGAAATCTGG	95
Db	46938	TGGCCACAGATGCTTTAAATCCAAAACCTGGCCGCTTGAGGCACAAAAGAAATCTGG	46997
OY	96	GTAATAATGATGTCCAAATCCATCGCCACCACTTAAATAGACGACACAACTAGTGAAGTGC	155
Db	46998	GTAATAATGATGTCCAAATCCATCGCCACCACTTAAATAGACGACACAACTAGTGAAGTGC	47057
OY	156	TGAGTAGGCTCTACAGAGTGAACGAGGAGTACACCCAAAACAAGAGGAGGACAGAGA	215
Db	47058	TGAGTAGGCTCTACAGAGTGAACGAGGAGTACACCCAAAACAAGAGGAGGACAGAGA	47117
OY	216	TCATCAAGAACCTCATCAAGACAGCATCAAGCTGGCCATTCTTATAGAAATTAATCACT	275
Db	47118	TCATCAAGAACCTCATCAAGACAGCATCAAGCTGGCCATTCTTATAGAAATTAATCACT	47177
OY	276	TTAATCAAGATGAGCTNGCATTTGATGAGAAATTTAAGAGAAAGTTCATCGCTTGGCTA	335
Db	47178	TTAATCAAGATGAGCTNGCATTTGATGAGAAATTTAAGAGAAAGTTCATCGCTTGGCTA	47237
OY	336	TGACCGTGATGACGTTTCATCAGGTGGATTAATACCTTTACCGGAATGTGTATCCAGGC	395
Db	47238	TGACCGTGATGACGTTTCATCAGGTGGATTAATACCTTTACCGGAATGTGTATCCAGGC	47297
OY	396	TGTTAAATGATGACAGAGAGATCTGCACCAATCATTCAGCGCCACTCACTGCCAAT	455
Db	47298	TGTTAAATGATGACAGAGAGATCTGCACCAATCATTCAGCGCCACTCACTGCCAAT	47358
OY	456	CACATGACCGGGTTAAATATGCTTTGATCATATTTTCAGATGTGAAATTTTGGCTGCT	515
Db	47358	CACATGACCGGGTTAAATATGCTTTGATCATATTTTCAGATGTGAAATTTTGGCTGCT	47417
OY	516	TGTATTAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGTGATGATATCAACA	575
Db	47418	TGTATTAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGTGATGATATCAACA	47477
OY	576	AAATGTGATGAAAGAAACATATGA 601	
Db	47478	AAATGTGATGAAAGAAACATATGA 47503	

```

US-10-029-386-22914
; Sequence 22914, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine, vers. 1.1
; SEQ ID NO 22914
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC035144.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EST_HUMAN HIT: AL549492.1, EVALU0 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P31390, EVALU0 2.60e-01
; OTHER INFORMATION: NT HIT: g17657123, EVALU0 0.00e+00
US-10-029-386-22914

```

Query Match 90.2%; Score 542.4; DB 16; Length 544;
Best Local Similarity 99.8%; Pred. No. 7.1e-152;

Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 36 TGCCACAGATGCTTTAATTCCTGAGCCGTTGAGGACCAAGAAAGATCTTGG 95
DB 1 TGCCACAGATGCTTTAATTCCTGAGCCGTTGAGGACCAAGAAAGATCTTGG 60
QY 96 GTTAAATGCTGCTCAATTCATGCGCCACCTTAAATAGACACAAAGTAGAGGTGC 155
DB 61 GTTAAATGCTGCTCAATTCATGCGCCACCTTAAATAGACACAAAGTAGAGGTGC 120
QY 156 TGAATGAGCTCTACAGAGTACCAAGGAGTACCCCAAGAAAGAGAGGAGAGAGA 215
DB 121 TGAATGAGCTCTACAGAGTACCAAGGAGTACCCCAAGAAAGAGAGGAGAGAGA 180
QY 216 TCATCAGAACTCATCAAGAGTATCAAGCTGCGCATCTTTATAGGAATATCAGT 275
DB 181 TCATCAGAACTCATCAAGAGTATCAAGCTGCGCATCTTTATAGGAATATCAGT 240
QY 276 TTAATCAAGATGAGTATGATGAGAAATTTAAGAAAGAAAGTTTATCAGTTGCTA 335
DB 241 TTAATCAAGATGAGTATGATGAGAAATTTAAGAAAGAAAGTTTATCAGTTGCTA 300
QY 336 TGAACGCTGCTCATTCATCAGTGTATATACCTTTGACCGGAATGTGTATCCAGC 395
DB 301 TGAACGCTGCTCATTCATCAGTGTATATACCTTTGACCGGAATGTGTATCCAGC 360
QY 396 TGTAAATGATGATGAGAGATGCTGACCAAAATCATTCAGCGCACCTTCACTGCGCAAT 455
DB 361 TGTAAATGATGATGAGAGATGCTGACCAAAATCATTCAGCGCACCTTCACTGCGCAAT 420
QY 456 CACATGAGCGGGTTAATATGTCTTGTATCATTTTTCAGATTTTGGTGGCT 515
DB 421 CACATGAGCGGGTTAATATGTCTTGTATCATTTTTCAGATTTTGGTGGCT 480
QY 516 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGATGATATACACA 575
DB 481 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGATGATATACACA 540
QY 576 AAAT 579
DB 541 AAAT 544

```

```

RESULT 11
US-10-087-192-493
; Sequence 493, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morrie, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 62231
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(62231)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-493

```

Query Match 74.2%; Score 446; DB 13; Length 62231;
Best Local Similarity 86.1%; Pred. No. 1.1e-121;

Matches 494; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

QY 28 TCCTCCAGTGGCCACAGATGCTTTAATTCCTGAGCCGTTGAGGACCAAGAAAGAA 87
DB 50644 TTCTGAGTGGCTACAGATGCTTTAATTCCTGAGCCGTTGAGGACCAAGAAAGAA 50703
QY 88 GATCTTGGGTTAAATGCTGCTCAATTCATGCGCCACCTTAAATAGACACAAAGTAG 147
DB 50704 GATCTTGGGTTAAATGCTGCTCAATTCATGCGCCACCTTAAATAGACACAAAGTAG 50763
QY 148 TGAATGCTGATGATGAGTCTTACAGAGTACCAAGGAGTACCCCAAGAAAGAGAGC 207
DB 50764 TGAATGCTGATGATGAGTCTTACAGAGTACCAAGGAGTACCCCAAGAAAGAGAGC 50823
QY 208 AGAAGATCATCAAGAACCTTCAAGACAGTATCAAGCTGCGCATCTTTATAGGA 267
DB 50824 AGAAGATCATCAAGAACCTTCAAGACAGTATCAAGCTGCGCATCTTTATAGGA 50883
QY 268 TTAATCAATTAATCAAGATGAGTATGAGAAATTTAAGAAAGAAAGTTTATCA 327
DB 50884 CATCAATTAATCAAGATGAGTATGAGAAATTTAAGAAAGAAAGTTTATCA 50943
QY 328 GCTTGTATGACCGTGTATGATTTTCAATGATGATGATTTATACCTTTGACCGGAATGTGT 387
DB 50944 GCTTGTATGACCGTGTATGATTTTCAATGATGATGATTTATACCTTTGACCGGAATGTGT 51003
QY 388 ATCAAGCTGTTTAAATGATGAGAGATGCTGACCAAAATCATTCAGCGCACCTTCACTGCGCAAT 447
DB 51004 ATCAAGCTGTTTAAATGATGAGAGATGCTGACCAAAATCATTCAGCGCACCTTCACTGCGCAAT 51063
QY 448 TGCAGATGATGAGAGCGGTTAATATGTCTTGTATCATTTTTCAGATTTTGGTGGCT 507
DB 51064 TGCAGATGATGAGAGCGGTTAATATGTCTTGTATCATTTTTCAGATTTTGGTGGCT 51123
QY 508 GGTGCTGCTGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGATGATATACACA 567
DB 51124 GGTGCTGCTGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGATGATATACACA 51183
QY 568 TATCAACAAATGTTGATGATGAGAGAAATATGA 601
DB 51184 CATCAACAAATGTTGATGATGAGAGAAATATGA 51217

```

RESULT 12
US-10-087-192-494
; Sequence 494, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 494
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-494

Query Match 73.9%; Score 444.4; DB 13; Length 2087;
Best Local Similarity 86.6%; Pred. No. 4.2e-122;
Matches 490; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 36 TGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTTCCAGGACAAAGAGATCTTG 95
DB 508 TGGCTACAGATGCTTTCAATCCAAAACCTGGCCGTTCCAGGACAAAGAGATCTTG 567
QY 96 GTAAATGGTGCTCAATTCATGCCACCACTTAAATAGACACAGATGAGAGTCG 155
DB 568 GCAAAATGGATCCAAATCCATGCGCACCGCTGATCGACACACAGAGAGAGTCG 627
QY 156 TGGATAGCTCTACAGATGACAGAGAGAGACCCAAACAGAAAGAGAGAGAGAG 215
DB 628 TGGATAGCTCTACAGAGAGAGAGACCCAAACAGAAAGAGAGAGAGAGAGAG 687
QY 216 TCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGGAATATCAGT 275
DB 688 TCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATCTTTATAGGAATATCAGT 747
QY 276 TTAATCAAGATGAGCTAGATGATGAGAGAAATTTAAGAAAGATTCACTGCTGCTA 335
DB 748 TCAATCAAGACGAGCTGCGCTCATGAGAGAAATTTCAAGAAAGAGTGCACAGCTTGCCA 807
QY 336 TGAACGATGCTTCCATCAGTGGATTAATCCTTTGACCGGAATGCTTATCCAGGC 395
DB 808 TGAACGATGCTTCCATCAGTGGATTAATCCTTTGACCGGAATGCTTATCCAGGC 867
QY 396 TGTAAATGATGACAGAGATGCTGCACCAATCAATTCAGCGCCCACTCATGCGCAAGT 455
DB 868 TGTGTAACAGATGCGGAGAGCTCTTACAGAGATCAATTCAGCGCCCACTCATGCGCAAGT 927
QY 456 CACATGAGAGGGTTAATATGCTTTGATCATTTTTCAGATGTAATTTTGGCTGCT 515
DB 928 CTCACGAGAGGGTTAATATGCTTTGATCATTTTTCAGATGTAATTTTGGCTGCT 987
QY 516 TGTATATGCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 575
DB 988 TGTACATATCTTTGGAATTTTAAACCTCATCTTACGAAACCTTTGGAGAGGATCAACA 1047
QY 576 AAATGTTGATGAGAGAAACATATGA 601
DB 1048 AAATGTTGATGAGAGAAACATATGA 1073

RESULT 13
US-09-796-692-3207

; Sequence 3207, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3207
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-3207

Query Match 47.5%; Score 285.4; DB 9; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 252 CCATCTTTATAGATTAATCAAGTTTAAATCAAGATGAGCTAGCATGATGAGAAATTTA 311
DB 1 CCATCTTTATAGATTAATCAAGTTTAAATCAAGATGAGCTAGCATGATGAGAAATTTA 60
QY 312 AGAAGAAATTTATCACTGCTTATGACCGTGTCACTTTCCATCAGGTGATTAATCCT 371
DB 61 AGAAGAAATTTATCACTGCTTATGACCGTGTCACTTTCCATCAGGTGATTAATCCT 120
QY 372 TTGACCGAATGTTATTCAGGCTGTTAAATGAATGAGAGATGCTGCACCAATCA 431
DB 121 TTGACCGAATGTTATTCAGGCTGTTAAATGAATGAGAGATGCTGCACCAATCA 180
QY 432 TTCAGGCACTCACTGCAAGTCAATGACAGGAGTAAATATGCTTTGATCATTTT 491
DB 181 TTCAGGCACTCACTGCAAGTCAATGACAGGAGTAAATATGCTTTGATCATTTT 240
QY 492 CAGATTGTAATTTTGGCTGCTGCTTATTAATCCTTTTGGGAATTT 538
DB 241 CAGATTGTAATTTTGGCTGCTGCTTATTAATCCTTTTGGGAATTT 287

RESULT 14
US-10-040-862-3207
; Sequence 3207, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Reiter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3207
LENGTH: 287
TYPE: DNA
ORGANISM: Homo sapiens
US-10-040-862-3207

Query Match 47.5%; Score 285.4; DB 14; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 252 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTAGAGTTCGATGAGAAATTTA 311
DB 1 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTAGAGTTCGATGAGAAATTTA 60
QY 312 AGAAGAAAGTTATCATCGCTGCTATGACCGGTGCTGTTCCATCAGGTGGATTATACCT 371
DB 61 AGAAGAAAGTTATCATCGCTGCTATGACCGGTGCTGTTCCATCAGGTGGATTATACCT 120
QY 372 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGAGTGTGCACCAATCA 431
DB 121 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGAGTGTGCACCAATCA 180
QY 432 TTGACCGCACTCATCTGCAAGTCAATGACAGGTTAATATGCTTTGATCAATTTT 491
DB 181 TTGACCGCACTCATCTGCAAGTCAATGACAGGTTAATATGCTTTGATCAATTTT 240
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 538
DB 241 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 287

RESULT 15
US-10-057-475B-3207
; Sequence 3207, Application US/10057475B
; Publication No. US20040002068A1

GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Ai-jun
APPLICANT: Ordenez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3207
LENGTH: 287
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-475B-3207

Query Match 47.5%; Score 285.4; DB 17; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 252 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTAGAGTTCGATGAGAAATTTA 311
DB 1 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTAGAGTTCGATGAGAAATTTA 60
QY 312 AGAAGAAAGTTATCATCGCTGCTATGACCGGTGCTGTTCCATCAGGTGGATTATACCT 371
DB 61 AGAAGAAAGTTATCATCGCTGCTATGACCGGTGCTGTTCCATCAGGTGGATTATACCT 120
QY 372 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGAGTGTGCACCAATCA 431
DB 121 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGAGTGTGCACCAATCA 180
QY 432 TTGACCGCACTCATCTGCAAGTCAATGACAGGTTAATATGCTTTGATCAATTTT 491
DB 181 TTGACCGCACTCATCTGCAAGTCAATGACAGGTTAATATGCTTTGATCAATTTT 240
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 538
DB 241 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 287

RESULT 16
US-10-154-884B-3207
; Sequence 3207, Application US/10154884B
; Publication No. US20040005561A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannon, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-154-884B-3207

Query Match      47.5%; Score 285.4; DB 17; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      252 CCATTCTTTATGGAATATCAAGTTTAATCAAGATGAGATGAGATGAGAAATTTA 311
DB      1 CCATTCTTTATGGAATATCAAGTTTAATCAAGATGAGATGAGATGAGAAATTTA 60
QY      312 AGAAGAAAGTTATCATGAGCTTGCTATGACCGTGTGCTGATTCCATCAGGTGATTTATACCT 371
DB      61 AGAAGAAAGTTATCATGAGCTTGCTATGACCGTGTGCTGATTCCATCAGGTGATTTATACCT 120
QY      372 TTGACCGGAATGTATATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB      121 TTGACCGGAATGTATATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 180
QY      432 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATGCTTTGATCATTTT 491
DB      181 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATGCTTTGATCATTTT 240
QY      492 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB      241 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 287

RESULT 17
US-10-764-324-3207
/ Sequence 3207, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannon, Jane
```

```
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ CURRENT FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-764-324-3207

Query Match      47.5%; Score 285.4; DB 19; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      252 CCATTCTTTATGGAATATCAAGTTTAATCAAGATGAGATGAGATGAGAAATTTA 311
DB      1 CCATTCTTTATGGAATATCAAGTTTAATCAAGATGAGATGAGATGAGAAATTTA 60
QY      312 AGAAGAAAGTTATCATGAGCTTGCTATGACCGTGTGCTGATTCCATCAGGTGATTTATACCT 371
DB      61 AGAAGAAAGTTATCATGAGCTTGCTATGACCGTGTGCTGATTCCATCAGGTGATTTATACCT 120
QY      372 TTGACCGGAATGTATATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB      121 TTGACCGGAATGTATATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 180
QY      432 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATGCTTTGATCATTTT 491
DB      181 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATGCTTTGATCATTTT 240
QY      492 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB      241 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 287

RESULT 18
US-10-796-692-2801/C
/ Sequence 2801, Application US/09796692
/ Publication No. US20020198362A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannon, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
/ TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
/ FILE REFERENCE: 2077.001200
/ CURRENT APPLICATION NUMBER: US/09/796,692
```

```

; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-2801
```

```

Query Match      47.2%; Score 283.8; DB 9; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 252 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTA 311
DB 287 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTTATCATGCTTGTATGACCGTGTGCTGATTCATCCATGAGTGATTAACCT 371
DB 227 AGAAGAAAGTTTATCATGCTTGTATGACCGTGTGCTGATTCATCCATGAGTGATTAACCT 168
QY 372 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTGACCGCACTCATCTGCCAAGTCACATGACGAGGTTAATATGCTTTGATCATTTT 491
DB 107 TTGACCGCACTCATCTGCCAAGTCACATGACGAGGTTAATATGCTTTGATCATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 1
```

```

RESULT 19
US-10-040-862-2801/c
; Sequence 2801, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
```

```

; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-2801
```

```

Query Match      47.2%; Score 283.8; DB 14; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 252 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTA 311
DB 287 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTTATCATGCTTGTATGACCGTGTGCTGATTCATCCATGAGTGATTAACCT 371
DB 227 AGAAGAAAGTTTATCATGCTTGTATGACCGTGTGCTGATTCATCCATGAGTGATTAACCT 168
QY 372 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTGACCGCACTCATCTGCCAAGTCACATGACGAGGTTAATATGCTTTGATCATTTT 491
DB 107 TTGACCGCACTCATCTGCCAAGTCACATGACGAGGTTAATATGCTTTGATCATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 1
```

```

RESULT 20
US-10-057-475B-2801/c
; Sequence 2801, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
```

```
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2801
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-2801
```

```
Query Match          47.2%; Score 283.8; DB 17; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 252 CCATTCTTATAGGAATATCAGTTTAATCAAGATGAGCATGATGAGAAATTTA 311
DB 287 CCATTCTTATAGGAATATCAGTTTAATCAAGATGAGCATGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTATCAGCTTGCTATGACCGTGTGCTCATGCTTCCATCAGTGATTAACCT 371
DB 227 AGAAGAAAGTTATCAGCTTGCTATGACCGTGTGCTCATGCTTCCATCAGTGATTAACCT 168
QY 372 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTCAGGCGCACTCCTGCAAGTCAATGACGCGTTAATATGCTTTGATCATTTT 491
DB 107 TTCAGGCGCACTCCTGCAAGTCAATGACGCGTTAATATGCTTTGATCATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 1
```

```
RESULT 21
US-10-154-884B-2801/C
/ Sequence 2801, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
```

```
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2801
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-154-884B-2801
```

```
Query Match          47.2%; Score 283.8; DB 17; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 252 CCATTCTTATAGGAATATCAGTTTAATCAAGATGAGCATGATGAGAAATTTA 311
DB 287 CCATTCTTATAGGAATATCAGTTTAATCAAGATGAGCATGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTATCAGCTTGCTATGACCGTGTGCTCATGCTTCCATCAGTGATTAACCT 371
DB 227 AGAAGAAAGTTATCAGCTTGCTATGACCGTGTGCTCATGCTTCCATCAGTGATTAACCT 168
QY 372 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTCAGGCGCACTCCTGCAAGTCAATGACGCGTTAATATGCTTTGATCATTTT 491
DB 107 TTCAGGCGCACTCCTGCAAGTCAATGACGCGTTAATATGCTTTGATCATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 1
```

```
RESULT 22
US-10-764-324-2801/C
/ Sequence 2801, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ CURRENT FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
```

```

; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-2801

Query Match      47.2%; Score 283.8; DB 19; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 252 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTACGCTTATGAGAAATTTA 311
DB 287 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTACGCTTATGAGAAATTTA 228

QY 312 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGATGATTAATCCT 371
DB 227 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGATGATTAATCCT 168

QY 372 TTGACCGGAATGTGTATTCAGGCTGTTAAATGATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTGTATTCAGGCTGTTAAATGATGACAGAGATGCTGCACCAATCA 108

QY 432 TTGACCGGCACTCATCTGCTCAAGTCAATGACAGCGGTTAATGCTCTTTGATCATTTT 491
DB 107 TTGACCGGCACTCATCTGCTCAAGTCAATGACAGCGGTTAATGCTCTTTGATCATTTT 48

QY 492 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 1

RESULT 23
US-09-796-692-5724
; Sequence 5724, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
```

```

; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (58)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (59)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-5724

Query Match      45.3%; Score 272.4; DB 9; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTACGCTTATGAGAAATTT- 310
DB 1 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTACGCTTATGAGAAATTTNC 60

QY 311 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGATGATTAATC 370
DB 61 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGATGATTAATC 120

QY 371 TTGACCGGAATGTGTATTCAGGCTGTTAAATGATGACAGAGATGCTGCACCAATC 430
DB 121 TTGACCGGAATGTGTATTCAGGCTGTTAAATGATGACAGAGATGCTGCACCAATC 180

QY 431 ATTACGGGCACTCATCTGCTCAAGTCAATGACAGCGGTTAATGCTCTTTGATCATTTT 490
DB 181 ATTACGGGCACTCATCTGCTCAAGTCAATGACAGCGGTTAATGCTCTTTGATCATTTT 240

QY 491 TCAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 538
DB 241 TCAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 288

RESULT 24
US-10-040-862-5724
; Sequence 5724, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01352005
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
```

```

; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (58)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (59)
; OTHER INFORMATION: n=A,T,C or G
; US-10-040-862-5724

```

```

Query Match      45.3%; Score 272.4; DB 14; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATTCTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATTT- 310
Db 1 CCATTCTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATNNC 60

QY 311 AAGAGAAAGTTTATAGCTTGTGCTTATGACCGTGTCAAGTTTCCATGAGTGAATYACC 370
Db 61 AAGAGAAAGTTTATAGCTTGTGCTTATGACCGTGTCAAGTTTCCATGAGTGAATYACC 120

QY 371 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATC 430
Db 121 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATC 180

QY 431 ATTACGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTTGATCATTTT 490
Db 181 ATTACGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTTGATCATTTT 240

QY 491 TCAGATGTGAATTTTGGCTGCTTGTAATATCCTTTTGGGAATTTT 538
Db 241 TCAGATGTGAATTTTGGCTGCTTGTAATATCCTTTTGGGAATTTT 288

```

```

RESULT 25
US-10-057-475B-5724
; Sequence 5724, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mammion, Jane
; APPLICANT: Clapper, Jonathan David

```

```

; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carrez, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(288)
; OTHER INFORMATION: n = G, A, C or T
; US-10-057-475B-5724

```

```

Query Match      45.3%; Score 272.4; DB 17; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATTCTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATTT- 310
Db 1 CCATTCTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATNNC 60

QY 311 AAGAGAAAGTTTATAGCTTGTGCTTATGACCGTGTCAAGTTTCCATGAGTGAATYACC 370
Db 61 AAGAGAAAGTTTATAGCTTGTGCTTATGACCGTGTCAAGTTTCCATGAGTGAATYACC 120

QY 371 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATC 430
Db 121 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATC 180

QY 431 ATTACGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTTGATCATTTT 490
Db 181 ATTACGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTTGATCATTTT 240

QY 491 TCAGATGTGAATTTTGGCTGCTTGTAATATCCTTTTGGGAATTTT 538
Db 241 TCAGATGTGAATTTTGGCTGCTTGTAATATCCTTTTGGGAATTTT 288

```

```

RESULT 26
US-10-154-884B-5724
; Sequence 5724, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(288)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-5724

Query Match      45.3%; Score 272.4; DB 17; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      252 CCATCTTTTAGGAATATCAAGTTTAATCAAGATGAGCTAGCATTTGATGAGAAATTT- 310
DB      1 CCATCTTTTAGGAATATCAAGTTTAATCAAGATGAGCTAGCATTTGATGAGAAATNNC 60

QY      311 AAGAAGAAAGTTATCATGAGTTGCTATGACCGTGTCAGTTTCATCAGTGATTTATAC 370
DB      61 AAGAAGAAAGTTATCATGAGTTGCTATGACCGTGTCAGTTTCATCAGTGATTTATACC 120

QY      371 TTTGACCGGAATGTGTTATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATC 430
DB      121 TTTGACCGGAATGTGTTATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATC 180

QY      431 ATTACGGGCACTCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 490
DB      181 ATTACGGGCACTCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 240

QY      491 TCAGATTGGAATTTTGGCTGCTGTATATCTTTTGGGAATTTT 538
DB      241 TCAGATTGGAATTTTGGCTGCTGTATATCTTTTGGGAATTTT 288

RESULT 27
US-10-764-324-5724
; Sequence 5724, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (58)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (59)
; OTHER INFORMATION: n=A,T,C or G
US-10-764-324-5724

Query Match      45.3%; Score 272.4; DB 19; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      252 CCATCTTTTAGGAATATCAAGTTTAATCAAGATGAGCTAGCATTTGATGAGAAATTT- 310
DB      1 CCATCTTTTAGGAATATCAAGTTTAATCAAGATGAGCTAGCATTTGATGAGAAATNNC 60

QY      311 AAGAAGAAAGTTATCATGAGTTGCTATGACCGTGTCAGTTTCATCAGTGATTTATAC 370
DB      61 AAGAAGAAAGTTATCATGAGTTGCTATGACCGTGTCAGTTTCATCAGTGATTTATACC 120

QY      371 TTTGACCGGAATGTGTTATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATC 430
DB      121 TTTGACCGGAATGTGTTATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATC 180

QY      431 ATTACGGGCACTCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 490
DB      181 ATTACGGGCACTCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 240

QY      491 TCAGATTGGAATTTTGGCTGCTGTATATCTTTTGGGAATTTT 538
DB      241 TCAGATTGGAATTTTGGCTGCTGTATATCTTTTGGGAATTTT 288

RESULT 28
```

```

US-10-029-386-25411
; Sequence 25411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25411
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012678.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.51
; OTHER INFORMATION: SWISSPROT HIT: Q61768, EVALUE 1.10e-01
; OTHER INFORMATION: NT HIT: APL2095.1, EVALUE 1.30e-01
; OTHER INFORMATION: EST_HUMAN HIT: BF744133.1, EVALUE 2.00e-59
US-10-029-386-25411

Query Match      39.5%; Score 237.4; DB 16; Length 527;
Best Local Similarity 65.7%; Pred. No. 2.5e-60;
Matches 346; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 75 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 134
DB 1 AGCCCAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 60
QY 135 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 194
DB 61 ATGACACCGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 120
QY 195 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 254
DB 121 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 180
QY 255 TTCTTATAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 314
DB 181 TCTCTACCGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 240
QY 315 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 374
DB 241 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 300
QY 375 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 434
DB 301 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 360
QY 435 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 494
DB 361 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAG 420
QY 495 ATTGTAATTTTGGTTCCTTGATTAATCTTTTGGGAAATTTTAAACCCCACTTAACA 554
DB 421 ATTGTAATTTTGGTTCCTTGATTAATCTTTTGGGAAATTTTAAACCCCACTTAACA 480
QY 555 AACTATGTATGTATCAACAAATGTGTGATGAAGAAACATATGA 601
DB 481 GATTTGTGAAGAAATCAATGAATGTTCTAGATGAAGAAAGTCTTTAA 527

RESULT 29
US-10-094-466-41
; Sequence 41, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Spylek et al.

```

```

; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 41
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26) ..(584)
US-10-094-466-41

Query Match      34.6%; Score 207.8; DB 17; Length 645;
Best Local Similarity 59.9%; Pred. No. 2.2e-51;
Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCCCTCCGATGGCCACAGATGCTTAAATCCAAAACCTGACGTCACACA 82
DB 8 CTGTGTCCCTCCGATGGCCACAGATGCTTAAATCCAAAACCTGACGTCACACA 67
QY 83 AAGAAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAGAGACACA 142
DB 68 AAGAAGATCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAGAGACACA 127
QY 143 AGTATGAGAGTCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAGAGACACA 202
DB 128 AGTATGAGAGTCTTGAGTAATGGTTCCTCAATCCATGCGCACCCTTAATAGAGACACA 187
QY 203 GAGGAGAGAGATCAAGAACTCATAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 262
DB 188 GAGGAGAGAGATCAAGAACTCATAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 247
QY 263 AGGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 322
DB 248 CGTGGGAGACAGCTGGGCGGTGAGAGAGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 307
QY 323 CATCAGCTTGATGACGCTGCTGATGCTTCAATGCTGATGCTTCAATGCTGATGCTTCAAT 382
DB 308 CCGTGGGAGACAGCTGGGCGGTGAGAGAGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 367
QY 383 GTTATCCAGGCTGTTAAATGAATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 442
DB 368 GTTATCCAGGCTGTTAAATGAATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 427
QY 443 CTCACCTGCAAGTCAATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 502
DB 428 CTCACCTGCAAGTCAATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 487

```


QY 503 TTTTGGCTGCTGTATATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGT 562
Db 488 TTCTGTGCTGCGCTCTACGGCCCGCCGAGCCCTACCGCTCCCACTGCGCAGGATCTGC 547
QY 563 GATGATATCAAAAATGTTGGATGAAGAGACATATGA 601
Db 548 GAGGGCTGGGCGGATGCTGAGACGAGGGCAGCTCTGA 586
RESULT 30
US-10-416-314-70
; Sequence 70, Application US/10416314
; Publication No. US20040082508A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: SANJANWALA, Madhusudan M.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: LU, Yan
; APPLICANT: GLEITZEN, Kimberly J.
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: HAPALIA, April J.A.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: WARREN, Bridget A.
; APPLICANT: HONCHELL, Cynthia D.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: LEE, Sally
; APPLICANT: XU, Yuming
; APPLICANT: YANG, Junning
; APPLICANT: IAL, Preeti G.
; APPLICANT: TRAN, Bao
; APPLICANT: ISON, Craig H.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: KAREHT, Stephanie K.
; TITLE OF INVENTION: SECRETED PROTEINS
; FILE REFERENCE: PI-0287 USN
; CURRENT APPLICATION NUMBER: US/10/416,314
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/247,505
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: US 60/249,642
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,824
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,824
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/254,305
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/256,448
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PERL Program
; SEQ ID NO 70
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No: 2013095CBI
US-10-416-314-70
Query Match 34.6%; Score 207.8; DB 18; Length 2108;
Best Local Similarity 59.9%; Pred. No. 4.4e-51;
Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGCGCACAGATGCTTTAATTCAAAACTGGCCGTTGACGGCAAA 82
Db 412 CTGTGTCCCGCCGACGGCCATGACACCTTCAGACCAAGAGCTGGCTGACAGGCGCAG 471
QY 83 AAGAAGATCTTGGTAAATGCTGCCAATTCATGGCCACCACTTAATAGACGACAA 142
Db 472 AAGAAGCTCTGTAGTAAGATGAGGCTCCAAAGCGAGTGCGCTGCTGAGTGAAC 531
QY 143 AGTAGTAGGTGCTGATGAGCTCTACAGATGACGAGGAGTACACCCAAAACAAGAG 202
Db 532 AGCAGTAGAGTGTGTGATGAGCTGTACCGGCCACCAAGAGTTCACGCCACCGCAAG 591
QY 203 GAGGCAGAGAAAGATCATCAAGAACTTCATCAAGACATCAAGCTGGCAATCTTTAT 262
Db 592 GAGGCCCAAGAAAGATGCTCAAGAACTGTCAAGTGCGCCCTGAAGCTGGAGACTGTGCTG 651
QY 263 AGGAATATCATGTTTATCAAGTGAAGTGAAGTGAAGTGAAGAAATTTAAGAAAGTT 322
Db 652 CGTGGGACCAAGCTGGGCGGCTGAGAGCTGGCGCTGCGGCGCTTCCGCCACCGGCG 711
QY 323 CATCAGCTTGCTATGACCGGTGTCAGTTTCATCAGGTGATTAATCATTTGACCGGAAT 382
Db 712 CGCTGCTTGCCATGACGGCCGTCAAGCTTTCACCAAGTGAATTCACCTTCGACCGGCG 771
QY 383 GTGTTATCCAGGCTGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCAC 442
Db 772 GTGCTGCGCGCGCGGCTGCTGAGTGCGCGGACCTGTGACCAAGCCGCGTCCCGAC 831
QY 443 CTCACGTGCCAATGTCATGAGACGGGTTAATATGCTTTGATATTTTTCAGTTGTGA 502
Db 832 CTGACCGCAAGTCCCAAGCGCCGATCAACCAAGTGTGCGCCACCTAGCGGCTCGAC 891
QY 503 TTTTGGCTGCTGTATATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGT 562
Db 892 TTCTGTGCTGCGCTCTACGGCCCGCCGAGCCCTACCGCTCCACCTGCGCAGATCTGC 951
QY 563 GATGATATCAAAAATGTTGGATGAAGAGACATATGA 601
Db 952 GAGGGCTGGGCGGATGCTGAGCAGGGGAGCCTCTGA 990

RESULT 31
US-09-816-828-5
; Sequence 5, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/816,828
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 5
; LENGTH: 1188
; TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (174)..(731)
 US-09-816-828-5

Query Match 34.2%; Score 205.6; DB 9; Length 1188;
 Best Local Similarity 60.0%; Pred. No. 1.4e-50;
 Matches 343; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

```

QY 30 CTCGATGCGCCACAGATGCTTTTAATTCAGAAACCTGGCCGCTTACAGCACAAGAGA 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163 CTGCTAAGCCATGACACCTTGACCAAGAGCTGGCTTGCGACGAGAAAGAAC 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 90 TCTTGGTAAATAGTGTCGAATTCATGCGCCACCTTAATAGACGACAAAGTAGTG 149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 223 TCTGTAGTAAGATGGGCTCAAGGAGTGCTGGCTGCTGGTGAATGACACAGAGTG 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 150 AGGTGCTGATGAGCTCTACAGAGTACAGGAGGATACCCAAACAAAGAGAGGACG 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 AGGTGCTGATGAGCTCTACAGGAGGATACCCAAAGGATACAGGAGGAGGAGGAGG 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 210 AGAAGATCAAGAACTTCATCAAGACATCAAGCTGGCCATTTTATATAGAAATA 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 343 AGAAGATGCTCAAGAACTTCATCAAGATGAGCTGGCCATTTTATATAGAAATA 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 270 ATCAGTTTATCAAGATGAGCTGATGATGAGAAATTAAGAAAGTTTATCAGC 329
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 403 ACCAGTGGGCGGTGAGAGAGCTGGCTGCTGGGCGCTTCCGCCACGGGCGGCTGCC 462
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 330 TTGCTATGACCGTGTCACTTTTCATCAGGTGATTAATACCGGATGTGTAT 389
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 463 TGGCCATGACGGCGGTCAAGCTTCAACAGGTGATCTTCAACCGGCGGCTGCTGG 522
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 390 CCAAGCTGTTAAATGAATGACAGAGATGCTGACCAATCATTCAGCGCACCTCACTG 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 523 CCGTCGGGCTGTCTGAGTCCGCGACCTGTGACACGAGCGCTGCCACTGACCG 582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 450 CCAAGTCATGAGAGGGTTAATATGTCTTGAATATTTTTCAGATGTGATTTTGG 509
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 583 CCAAGTCCACGGCCGATCAACCAAGTGTCCGCCACCTAGCGAGCTGAGACTTCCGG 642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 510 CTGCTTGTATATCTTTTGGGAATTTAAACCCCACTTACAAAACTATGTATGTA 569
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 643 CTGCGCTTACAGGCCCCCGCGAGCCCTACCGCTCCACCTGCGCAGAGATGTGAGAGGCC 702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 570 TCAACAAATGTTGGATGAAAGACATATGA 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 703 TGGGCGGATGCTGAGACGAGGGCAGCCTCTGA 734
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 32
 US-10-094-466-43

; Sequence 43, Application US/10094466
 ; Publication No. US2003020363A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Spvtek et al
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
 ; TITLE OF INVENTION: AND METHODS OF USING
 ; FILE REFERENCE: 21402-290D
 ; CURRENT APPLICATION NUMBER: US/10/094,466
 ; PRIOR FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: 60/274,281
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 60/288,148
 ; PRIOR FILING DATE: 2001-05-02
 ; PRIOR APPLICATION NUMBER: 60/274,849
 ; PRIOR FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 60/275,235
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: 60/338,375
 ; PRIOR FILING DATE: 2001-12-04

```

; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 43
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(566)
US-10-094-466-43

```

Query Match 33.8%; Score 203.4; DB 17; Length 619;
 Best Local Similarity 60.0%; Pred. No. 4.4e-50;
 Matches 339; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

```

QY 37 GGCACAGATGCTTTAATTCAGAAACCTGGCCGCTTACAGCACAAGAGATCTGGG 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4 GGCATGACACCTTCAGACCAAGAGCTGCTCTGACGGGCGAGAAAGACTCTGTAG 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 97 TAAATGTGTGTCGAATCCATGCGCACCTTAATAGACGACAGATGATGAGTGCT 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 TAAAGTGGGTCCAAAGGAGAGTGCTGGCGGTGCTGTGATGACACAGAGTGAAGTGCT 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 GATAGACTCTACAGATGACCAAGGAGTACACCCAAACAAAGAGAGGACAGAAAT 216
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 124 GGATGAGCTGTACCGCGCACCAAGGAGTTCACGCGAGCCGCAAGAGGCGCAGAAAT 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 CATCAAGAACCTTCATCAAGACATGATCAAGCTGGCCATTTTATAGAAATATCAGT 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 GCTCAAGAACTGTGTCAAGGTGCGCTGAGAGCTGAGACTGTGCTGCGGGGACAGCT 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 TAATCAAGATGACTAGCATTTGATGAGAGAAATTTAAGAAAGTTTCACTTGTCTAT 336
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 GGGCGGTAGAGAGCTGGCGCTGTGCGGCCACCGGGGCGGCTGCTGGCAT 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 337 GACCGTGTCACTTTCATCAGGTGATTAACCTTTGACCGGAATGTGTATTCAGGCT 396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 304 GACGGCGGTCAAGCTTCCACAGGTGAACTTCACTTTCGACCGGCGGTGCTGGCGCGG 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 397 GTTAAATGAAATGACAGAGATGCTGACCAATCATTCAGCGGCCACTGACGCAAGTC 456
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 364 GCTGCTCAGTCCCGGAGCTGTCTGACACAGGCGGTGGGTCCCACTGACGCAAGTC 423
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 ACATGACCGGTAAATTAATGTCTTTGATCAATTTTCAAGTTTGGTGCCTT 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 424 CCAAGCGCCATCAACACAGTGTGGGCCCACTTACCGACATCTGCTGCTGGCT 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 GTATAATCTTTTGGGAATTTAAACCCCACTTACAAAACTATGTATGATATCAACA 576
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 484 CTACGGCCCCCGCGAGCCCTACCGCTCCACCTGCGCAGAGATCTGAGGGGCTGGGCGG 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 577 AATGTTGATGAAAGAACATATGA 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 544 GATGCTGACGAGGCGAGGCTCTGA 568
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 33

US-10-764-420-2116
 ; Sequence 2116, Application US/10764420
 ; Publication No. US20050084872A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lum, Pek Yee

```

; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; TITLE OF INVENTION: Possesses A Defined Biological Activity
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764.420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 80
; OTHER INFORMATION: n = A,T,C or G
;
US-10-764-420-2116

```

```

Query Match          33.2%; Score 199.8; DB 21; Length 1389;
Best Local Similarity 59.7%; Pred. No. 8.6e-49;
Matches 336; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY 39 CCACAGATGCTTTAATTCACAAAACCTGGCCGTTACAGCACAAGAGATCTTGGGTA 98
DB 82 CCATGGACACCTTACAGACGAGAGCCTGGCCCTGACAGGCCCAAGAAAGTCTTGCA 141
QY 99 AATGCTGTCCAAATCCATGCGCCACCACTTAATAGACACAAAGTAGAGGTGCTGG 158
DB 142 AGATGGCTTCCAGAGCCATGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 201
QY 159 ATGAGCTCTACAGAGTGACCAAGGAGTACACCCAAAACAAGAGAGGAGAGATCA 218
DB 202 ATGAGCTGTACAGAGCCACAGAGAGTTCACGCGCACGCGAGAGAGGAGGATG 261
QY 219 TCAAGAACCTATCAGACAGTCAATCAAGCTGCGCATTTCTTATAGGAATATCATTTA 278
DB 262 TGAAGAACCTGTGTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 321
QY 279 ATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGAAAGTTCATGAGTGTCTATGA 338
DB 322 ACAGCAATGAGCTGCGCCAGCTGACAGCGGTTCCGCGGCGCGGTCCGCAACTGCGCATGA 381
QY 339 CCGTGTGAGTTTCATCAGGTGATTAATCCTTTGACCGGAATGTGTATCCAGGCTGT 398
DB 382 CAGCCCTCAGCTTCCACAGGTAGATTTCACTTTGACCGCGCGTGTGTGTGTGTGTGTGTGT 441
QY 399 TAAATGATCAGAGAGATCTGACCAAAATCATTTAGCGCCACCTGACTGCGCAAGTCAAC 458
DB 442 TGTGTGAGTGCAGAGGACCTGTGTGACCAAGGTATGTGCGCGGACCTTACCTGCGCAAGTCCC 501
QY 459 ATGAGCGGGTTAATATGTCTTTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTGT 518
DB 502 ACCGCGCATCAATCAATCATCTTTCAGTCACTTTGCAATGTGATCTTCTGCGCGCGCTGT 561
QY 519 ATAATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGTATCAACAAA 578
DB 562 ACAGCCCAAGAGAGCCCTTACCGAGCATGTGTGCGCATGTGTGAGGCTGTGGAGGA 621
QY 579 TGTGTGATGAAGAGACATATGA 601
DB 622 TGTGTGACGAGGCTGCGCATCTGA 644

```

```

RESULT 34
US-09-822-849A-81
; Sequence 81, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulubeta, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2001-09-04
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Homo sapiens
;
US-09-822-849A-81

```

```

Query Match          30.3%; Score 182; DB 9; Length 1113;
Best Local Similarity 58.7%; Pred. No. 1.7e-43;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGAGTGGCCACAGATGCTTTAATTCACAAAACCTGGCCGTTACAGGACACAAAGAGA 89
DB 52 CTCGAGAGCCACATGAGAGTCTTTCATGCTCAAAAGCCCTGGACATGCAAGAGAGAGAGC 111
QY 90 TCTTGGTAAATGTGTGTCCAAATTCATGCGCCACCACTTAATAGACACAAAGTAGTG 149
DB 112 TACTGATGATAGTGGCGGCTGCTGTGTGCTCATCTTCTTATAGATGAGACAAAGCAGTG 171
QY 150 AGTGTCTGATGAGCTCTTACAGAGTACCCAGGAGTACACCCAAAACAAGAGAGCAG 209
DB 172 AGGTGTGATGATGAGCTTCAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 231
QY 210 AGAAGATCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGCGCATTTCTTATAGATA 269
DB 232 AGCGGCTGATCAAGAGACTGATCAAGTGGCCATCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 291
QY 270 ATCAGTTTAATCAAGATGAGCTAGCTTGTGATGAGAAATTTAAGAAAGAAAGTTCATGAGC 329
DB 292 GCTCTTTGGCCCAAGTAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 351
QY 330 TTGCTATGACCGGTGTGATGCTTTCATCAGGTGATTAATCCTTTGACCGGAATGTGTAT 389
DB 352 GTGCCATGAGCGGACCTTAGCTTTGTGTGAGGTAGATTTCACTTGTGAGGCTGTGTGTGT 411
QY 390 CCAAGCTGTAAATGATGAGAGAGATGCTGACCAAAATCATTTAGCGCCACCTGACTG 449
DB 412 CTGCGCTGCTGACCGAGTGTGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471
QY 450 CCAAGTCAATGAGCGGTTAATATGTCTTTGATCATTTTTCAGATGTGAATTTTGG 509
DB 472 CCAAGTCAATGAGCGGATGCGGACGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 531
QY 510 CTGCTTTGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGTGTA 569
DB 532 CGGCGCTCTATGTGGCTT-----GACCTTCACTGAGACCTTTGCAAGATCTGTGAGCGAC 585
QY 570 TCAACAAAATGTGTGATGAAGAGACATATGA 601
DB 586 TCAAGAAAGCTGTAGAGAGAGAGAGAGCTCTGA 617

```

```

RESULT 35
US-10-363-374-15
; Sequence 15, Application US/10363374
; Publication No. US20050048483A1
; GENERAL INFORMATION:

```

RESULT 36
US-10-302-172-537
; Sequence 537, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Driemanc, Radje T.

```

RESULT 37
US-10-959-539-59
; Sequence 59, Application US/10959539
; Publication No. US20050048623A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.

```

```

APPLICANT: HILLMAN, Jennifer L.
APPLICANT: LAL, Preeti
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: AU-YOUNG, Janice
APPLICANT: BANDMAN, Olga
APPLICANT: AZIMZAI, Yalda
APPLICANT: YANG, Junming
APPLICANT: LU, Dying Aina M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: PATTERSON, Chandra
APPLICANT: SHAH, Purni
TITLE OF INVENTION: CELL CYCLE AND PROLIFERATION PROTEINS
FILE REFERENCE: PR-0722 PCT
CURRENT APPLICATION NUMBER: US/10/959,539
PRIOR FILING DATE: 2004-10-06
PRIOR APPLICATION NUMBER: US/10/031,915
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: 60/145,075; 60/153,129; 60/164,647
PRIOR FILING DATE: 1999-07-21; 1999-09-08; 1999-11-10
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PERL Program
SEQ ID NO 59
LENGTH: 1268
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 1558289CB1
US-10-959-539-59

```

```

Query Match      30.3%; Score 182; DB 21; Length 1268;
Best Local Similarity 58.7%; Pred. No. 1.8e-43;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

```

```

QY 30 CTCGATGCGCAGATGCTTTAATTCGAAACCTGGCCCTTCAGGCAAGAAAGA 89
DB 92 CTCGAGACCCATGAGGCTCTTACGCTCAAGAGCCCTGACCTGCAAGCAAGAAAG 151
QY 90 TCTTGGGTAATAGTGCTCAATTCATCGCCACCACTTAATAGACACACAAGTAGT 149
DB 152 TACTGAGTAAGAGGGGGGCTCTGTGCTCATCTTTCATAGATGAGACAGAGT 211
QY 150 AGGTGCTGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAACAAAGAGAGG 209
DB 212 AGGTGCTGATGAGCTCTACAGAGTACCAAGGAGTACACAGCAGCCGCCAGGCC 271
QY 210 AGAAGTCAATCAAGAACCTCATCAAGACATCATCAAGCTGGCCATTTTATAGAA 269
DB 272 AGGCGCTGATCAAGGAGCTGATCAAGGCTCATCAAGGCTGTGCTGACCCGCA 331
QY 270 ATCAGTTAATCAAGATGAGCTGATGATGAGAAATTAAGAAAGTTATCATCAG 329
DB 332 GCTCCTTTGGCCCATGAGCTGCTGCTGCTACCCGCTTCCGCAAGCTCGGAG 391
QY 330 TTGCTATGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389
DB 392 GTGCGCATGACGAGCTTACCTTTGAGAGTACCTTCACTTCAAGGCTGCTGCT 451
QY 390 CCAAGCTGTTAATGATGAGAGATGCTGACCAAAATCATTCAGCCGCTCATCTG 449
DB 452 GTGCGCTGCTGACGAGGCTCGGAGAGTGTGTAGAGTTGGAGAACCACTCAG 511
QY 450 CCAAGTCAATGACGAGGCTTAATATGCTGTTGATGATTTTTCAGATTGGAAT 509
DB 512 CCAAGTCAATGACGAGGCTTAATATGCTGTTGATGATTTTTCAGATTGGAAT 571
QY 510 CTGCTGTTGATGATGCTTTTGGGAATTTTAAACCCCACTTACAAAACTATG 569
DB 572 CGGCTCTATGAGGCT-----GACTTCACTGAGACCTTGGCAAGATCTGTGAG 625
QY 570 TCAAGAAATGTTGATGAGAGACATATGA 601

```

```

DB 626 TCAGGAAGCTCTAGACGAAGGAGAGCTCTGA 657

```

```

RESULT 38
US-10-029-386-9211
Sequence 9211, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
PRIOR FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 9211
LENGTH: 500
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC035144.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
OTHER INFORMATION: NT HIT: g114725421, EVALUE = 2.4
OTHER INFORMATION: EST_HUMAN HIT: B6178783.1, EVALUE 0.00e+00
US-10-029-386-9211

```

```

Query Match      28.0%; Score 168.4; DB 16; Length 500;
Best Local Similarity 99.4%; Pred. No. 1.2e-39;
Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 432 TTCAGGCGCACCCTACTGCGCAATGCATGACGAGGTTAATATGCTTTGATCATTTT 491
DB 1 TTCAGGCGCACCCTACTGCGCAATGCATGACGAGGTTAATATGCTTTGATCATTTT 60
QY 492 CAGATTGTGAATTTTGGCTGCTTGTATATCTTTTGGGAATTTTAAACCCACTTAC 551
DB 61 CAGATTGTGAATTTTGGCTGCTTGTATATCTTTTGGGAATTTTAAACCCACTTAC 120
QY 552 AAAAATATGATGATGATCAACAAATGTTGATGAGAGACATATGA 601
DB 121 AAAAATATGATGATGATCAACAAATGTTGATGAGAGACATATGA 170

```

```

RESULT 39
US-09-918-995-35521

```

```

Sequence 35521, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: HySeq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 35521
LENGTH: 422
TYPE: DNA
ORGANISM: Homo sapiens
US-09-918-995-35521

```

```

Query Match      21.4%; Score 128.4; DB 10; Length 422;
Best Local Similarity 60.7%; Pred. No. 1.1e-27;

```

RESULT 41
US-10-191-803-330/c
; Sequence 330, Application US/10191803

RESULT 42
US-09-796-692-4917/c
; Sequence 4917, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077_001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950

```

; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (77)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-4917

Query Match          10.5%; Score 63; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTACAAAACATGATGATGATCAACAATAATGTTGATGAAGAACAATA 538
Db 538 AAACCCACTTACAAAACATGATGATGATGATCAACAATAATGTTGATGAAGAACAATA 479

QY 599 TGA 601
Db 478 TGA 476

RESULT 43
US-10-040-862-4917/C
; Sequence 4917, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
```

```

; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (77)
; OTHER INFORMATION: n=A,T,C or G
; US-10-040-862-4917

Query Match          10.5%; Score 63; DB 14; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTACAAAACATGATGATGATCAACAATAATGTTGATGAAGAACAATA 538
Db 538 AAACCCACTTACAAAACATGATGATGATGATCAACAATAATGTTGATGAAGAACAATA 479

QY 599 TGA 601
Db 478 TGA 476

RESULT 44
US-10-057-475B-4917/C
; Sequence 4917, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

/ LOCATION: (1)...(538)
/ OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-4917

Query Match 10.5%; Score 63; DB 17; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 598
DB 538 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 479

QY 599 TGA 601
DB 478 TGA 476

RESULT 45

US-10-154-884B-4917/c
/ Sequence 4917, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:

/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B

/ PRIOR FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO 4917
/ LENGTH: 538
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(538)
/ OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-4917

Query Match 10.5%; Score 63; DB 17; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 598
DB 538 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 479

QY 599 TGA 601
DB 478 TGA 476

Search completed: July 28, 2005, 19:13:31
Job time : 593 secs

25	545.4	90.7	580	5	BP302623	BP302623
26	547.4	90.2	595	5	CA9833219	irf55c03.1
27	538.6	89.6	834	5	BK405157	BK405157
28	538.4	89.6	747	6	CD520304	AGENCOUR
29	537	89.4	622	6	CD701806	CD701806
30	535	89.0	580	5	BP228279	BP228279
31	533.6	88.8	672	7	CN105328	17000599
32	533.4	88.8	566	7	CV023086	5463
33	530.6	88.3	707	7	CR557552	CR557552
34	529.8	88.2	572	7	CK903031	irf55c03.1
35	529.4	88.1	697	1	AV713586	AV713586
36	527.6	87.8	739	4	BG434554	BG434554
37	525.6	87.6	652	7	CR771358	CR771358
38	519.8	86.5	685	4	BG534704	60255386
39	519.2	86.4	593	7	CV027255	5464
40	513.6	85.5	757	6	CD523203	AGENCOUR
41	509.4	84.8	588	6	CD7233203	o19b07.1
42	497.8	82.8	583	5	BP345938	BP345938
43	496.6	82.6	605	7	CR771295	DKF2p4659
44	494.6	82.3	744	4	BI760235	60304517
45	489.2	81.4	773	5	BP433798	BP433798
46	486.6	81.0	616	5	CB545457	MMSP00711
47	486.2	80.9	640	6	CD639093	AGENCOUR
48	484.6	80.6	853	5	CD522223	AGENCOUR
49	481.8	80.2	582	5	BP261675	BP261675
50	479.6	79.8	588	7	CR787177	DKF2p4659
51	479.6	79.8	669	9	CN789950	41224483
52	476.4	79.3	681	6	CD663714	EST234. h
53	475.4	79.1	607	6	CB552439	CB552439
54	475	79.0	671	6	CB551156	MMSP0066
55	473.6	78.8	857	6	CB961433	AGENCOUR
56	470.2	78.2	771	6	CD521059	AGENCOUR
57	466.8	77.7	622	6	CB553311	MMSP0007
58	464.8	77.3	652	6	CD638280	AGENCOUR
59	464.4	77.3	582	5	BP379054	BP379054
60	456	75.9	942	5	CA984413	AGENCOUR
61	454.4	75.6	583	5	BP292302	BP292302
62	451.6	75.5	670	5	BP459271	BP459271
63	451.2	75.1	580	5	BP378379	BP378379
64	449.8	74.8	651	4	BG146795	60b95h09
65	449.6	74.8	908	4	BI668217	603b1471
66	446.4	74.3	664	6	CD639771	AGENCOUR
67	444.4	73.9	882	2	BF607236	MY1. 0001
68	444.4	73.9	852	4	BI525793	60282475
69	443.4	73.8	766	4	BI683693	60330624
70	442.8	73.7	841	6	CA469176	AGENCOUR
71	441.2	73.4	1542	3	AK090316	Mus
72	438	72.9	710	6	CD639143	AGENCOUR
73	437.2	72.7	800	4	BG534149	60255326
74	435.4	72.4	698	6	CD640827	AGENCOUR
75	434	72.2	743	6	CD639074	AGENCOUR
76	433.4	72.1	582	6	CD691893	EST84438
77	433.4	72.1	772	4	BI103517	AGENCOUR
78	433	72.0	656	6	CD691511	EST8034
79	431.4	71.8	656	4	BI687787	60331410
80	431.2	71.7	671	6	BY743355	BY743355
81	430.4	71.6	580	5	BP375383	Mus
82	430.4	71.6	688	4	BG146696	mad94h09
83	430.2	71.6	1836	3	AK087478	Mus
84	429.2	71.4	719	2	BF140519	mus
85	429.2	70.6	711	4	BG372935	60284060
86	418.4	69.4				

98	395.4	65.8	523	4	BM194176	TCAPAD1011
99	395	65.7	582	5	BP277360	BP277360
100	394.6	65.7	587	1	AU136721	AU136721

ALIGNMENTS

RESULT 1
LOCUS CF593784 806 bp mRNA linear EST 26-SEP-2003
DEFINITION AGENCOURT 15624053 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:30528465 5', mRNA sequence.
ACCESSION CF593784
VERSION CF593784.1 GI:36347693
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 806)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999).
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: NDAM614 row: h column: 10
High quality sequence stop: 648.
Location/Qualifiers

FEATURES

source

1..806
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30528465"
/feature_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_147"
/note="Organ: Placenta; Vector: pBluescriptR; Site 1:
all-Khoi; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to R05. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN

Query Match 97.9%; Score 588.4; DB 7; Length 806;
Best Local Similarity 99.7%; Pred. No. 3.9e-152;
Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY	1	CTTACGCGTCCGGCGCCGCTGC-CGACTCTCTCGATGGCCACAGATGCTTTAATTCOA	59
DB	103	CTTACGCGTCCGGCGCCGCTGCCTCTCCGATGGCCACAGATGCTTTAATTCOA	162
QY	60	AAACCTGGCCGCTTCAAGCAAAAGAGTCTGGTAAATGCTGCCAATTCATG	119
DB	163	AAACCTGGCCGCTTCAAGCAAAAGAGTCTGGTAAATGCTGCCAATTCATG	222
QY	120	CCACCACTTAATAGACGACACAGTAGTGGTGTGATGAGCTCTACAGAGTACCA	179

DB	223	CCACCACCTTAATAGACGACACAGTAGTGGTGTGATGAGCTCTACAGAGTACCA	282
QY	180	GGGACTTAACCCAAAAGAGGAGGAGAAATATCATGAAGACCTCATAGACAG	239
DB	283	GGGAGTACCCAAAAGAGGAGGAGAAATATCATGAAGACCTCATAGACAG	342
QY	240	TCATCAAGCTGGCATTCTTTATAGAAATATCAATTAAATCAAGATGAGTTCGA	299
DB	343	TCATCAAGCTGGCATTCTTTATAGAAATATCAATTAAATCAAGATGAGTTCGA	402
QY	300	TGAGAAATTTAAGAAAGTTCATCAGCTTGTATGACCGTGTGATTCATCAGG	359
DB	403	TGAGAAATTTAAGAAAGTTCATCAGCTTGTATGACCGTGTGATTCATCAGG	462
QY	360	TGGATTATACCTTTACCGGAATGTGTATCCAGGCTGTAAATGAAGCAGAGATGC	419
DB	463	TGGATTATACCTTTACCGGAATGTGTATCCAGGCTGTAAATGAAGCAGAGATGC	522
QY	420	TGCACCAATCATTTAGCGCCACCTCAGTGCACATGACATGACGGGTTAATGCT	479
DB	523	TGCACCAATCATTTAGCGCCACCTCAGTGCACATGACATGACGGGTTAATGCT	582
QY	480	TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGAAATTTTA	539
DB	583	TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGAAATTTTA	642
QY	540	AACCCCACTTAATAAATCATGTGATGTATCAAAATGTGTGAAGAAGACATAT	599
DB	643	AACCCCACTTAATAAATCATGTGATGTATCAAAATGTGTGAAGAAGACATAT	702
QY	600	GA 601	
DB	703	GA 704	

RESULT 2

CF272384

LOCUS

DEFINITION

AGENCOURT 15178772 NIH MGC 192 Homo sapiens cDNA clone

IMAGE:30512832 5', mRNA sequence.

CF272384

VERSION

CF272384.1 GI:33628296

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 920)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Agencourt

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

<http://image.llnl.gov>

Plate: NDAM573 row: m column: 01

High quality sequence stop: 598.

Location/Qualifiers

1..920

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30512832"

/feature_type="Brain glioblastoma"

/lab_host="DH10B (T1 phage-resistant)"

/clone lib="NIH MGC 192"
 /note="Vector: pExpress1; Site 1: SmaI; Site 2: NotI; The library was constructed by reverse transcription of 1 ug mRNA using the oligo dt primer GCGGCGGCC(1)20 and an RNaseH + MMLV reverse transcriptase. Second strand synthesis was carried out by standard methods. The cDNA was size selected by agarose gel for > 1.2 kb, digested with Not I and directionally cloned into the vector Express-1 at the SmaI/NotI sites. DNA from the primary library was used for in vitro transcription from the T7 promoter to produce biotinylated RNA transcripts. These biotinylated transcripts, along with blocking oligos to the poly-A, multiple cloning site and primer regions, were hybridized with single stranded circles produced by phageid production from the primary library to a Cot value of 10-20. Streptavidin/phenol extraction was utilized to remove DNA:RNA hybrids leaving un-hybridized single stranded circles which were repaired by primer extension and transformed back into E. coli resulting in the normalized library. Average insert size 2.0 kb. 3 linker/adaptor sequence GCGGCGGCC(1)20. This library was constructed by Agencourt Bioscience."

ORIGIN

```

Query Match      97.9%; Score 588.4; DB 7; Length 920;
Best Local Similarity 99.7%; Pred. No. 4.1e-152;
Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTTGAGCGTCCGCGCGCGTGC-CGACTCTCCGATGGCCAGATGCTTTAATCCA 59
    |||
Db 63 CTTGAGCGTCCGCGCGCGTGC-CGACTCTCCGATGGCCAGATGCTTTAATCCA 122

QY 60 AAAAAGCTGGCGTTCAGGCAAAAAGATCTTGGTAAATGCTGCCAATCCATCG 119
    |||
Db 123 AAAAGCTGGCGTTCAGGCAAAAAGATCTTGGTAAATGCTGCCAATCCATCG 182

QY 120 CCACCACTTAATAGACAGACAAAGTAGTAGGTGCTGATGAGCTTACAGAGTACCA 179
    |||
Db 183 CCACCACTTAATAGACAGACAAAGTAGTAGGTGCTGATGAGCTTACAGAGTACCA 242

QY 180 GGGAGTACACCCAAAACAAGAGGAGAGAGATCATCAAGAACTCATCAAGACAG 239
    |||
Db 243 GGGAGTACACCCAAAACAAGAGGAGAGAGATCATCAAGAACTCATCAAGACAG 302

QY 240 TCATCAAGCTGGCCATTCTTTATAGAAATATCATGTTAATCAAGATGACATGTA 299
    |||
Db 303 TCATCAAGCTGGCCATTCTTTATAGAAATATCATGTTAATCAAGATGACATGTA 362

QY 300 TGAAGAAATTTAAGAAAGTTTCATGCTGTATGACCGGTGTCAGTTCCATCAGG 359
    |||
Db 363 TGAAGAAATTTAAGAAAGTTTCATGCTGTATGACCGGTGTCAGTTCCATCAGG 422

QY 360 TGAATTATACCTTTGACCGGAATGTGTATCCAGGCGTTAAATGAATGAGAGATGC 419
    |||
Db 423 TGAATTATACCTTTGACCGGAATGTGTATCCAGGCGTTAAATGAATGAGAGATGC 482

QY 420 TGAACCAATCATTCAGCGCCACCTCATGCGCAAGTCACTGACGGGTTAATATGCT 479
    |||
Db 483 TGAACCAATCATTCAGCGCCACCTCATGCGCAAGTCACTGACGGGTTAATATGCT 542

QY 480 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTA 539
    |||
Db 543 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTA 602

QY 540 AACCCCACTTACAAAACCTATGTGATCAACAAAATGTTGATGAGAGAACTAT 599
    |||
Db 603 AACCCCACTTACAAAACCTATGTGATCAACAAAATGTTGATGAGAGAACTAT 662

QY 600 GA 601
    ||
Db 663 GA 664

```

RESULT 3

CN305329

LOCUS CN305329

DEFINITION 1700600027114 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN305329

VERSION CN305329.1 GI:47321743

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Branderberger R., Wei H., Zhang S., Lei S., Murge J., Fisk G., J., Li Y., Xu C., Pang R., Guegler K., Rao M. S., Mandalam R., Lebkowski J. and Stanton L. W.

TITLE Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6) 707-716 (2004)

COMMENT Contact: Branderberger R

Regenerative Medicine

Geron Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbranderberger@geron.com

Insert Length: 737 Std Error: 0.00.

Location/Qualifiers

1. 737

/organism="Homo sapiens"

/mol_type="RNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"

/clone lib="GRN PREHEP"

/note="oligo dt primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN

```

Query Match      97.6%; Score 586.8; DB 7; Length 737;
Best Local Similarity 99.5%; Pred. No. 1.1e-151;
Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CTTGAGCGTCCGCGCGCGTGC-CGACTCTCCGATGGCCAGATGCTTTAATCCA 59
    |||
Db 72 CTTGAGCGTCCGCGCGCGTGC-CGACTCTCCGATGGCCAGATGCTTTAATCCA 131

QY 60 AAAAAGCTGGCGTTCAGGCAAAAAGATCTTGGTAAATGCTGCCAATCCATCG 119
    |||
Db 132 AAAAAGCTGGCGTTCAGGCAAAAAGATCTTGGTAAATGCTGCCAATCCATCG 191

QY 120 CCACCACTTAATAGACAGACAAAGTAGTAGGTGCTGATGAGCTTACAGAGTACCA 179
    |||
Db 192 CCACCACTTAATAGACAGACAAAGTAGTAGGTGCTGATGAGCTTACAGAGTACCA 251

QY 180 GGGAGTACACCCAAAACAAGAGGAGAGAGATCATCAAGAACTCATCAAGACAG 239
    |||
Db 252 GGGAGTACACCCAAAACAAGAGGAGAGAGATCATCAAGAACTCATCAAGACAG 311

QY 240 TCATCAAGCTGGCCATTCTTTATAGAAATATCATGTTAATCAAGATGACATGTA 299
    |||
Db 312 TCATCAAGCTGGCCATTCTTTATAGAAATATCATGTTAATCAAGATGACATGTA 371

QY 300 TGAAGAAATTTAAGAAAGTTTCATGCTGTATGACCGGTGTCAGTTCCATCAGG 359
    |||
Db 372 TGAAGAAATTTAAGAAAGTTTCATGCTGTATGACCGGTGTCAGTTCCATCAGG 431

QY 360 TGAATTATACCTTTGACCGGAATGTGTATCCAGGCGTTAAATGAATGAGAGATGC 419
    |||
Db 432 TGAATTATACCTTTGACCGGAATGTGTATCCAGGCGTTAAATGAATGAGAGATGC 491

QY 420 TGAACCAATCATTCAGCGCCACCTCATGCGCAAGTCACTGACGGGTTAATATGCT 479
    |||

```

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Dn	162	AAAACTGGCGCTTACGAGCCACAAAAAGAAAGATCTTTGGGTAAATAATGATGTCCAAATTCATCTCG	221					
Qy	120	CCACCACTTAAATAGACGACACAAATGATGATGCTGGATGAGCTCTTACAGATGACCA	179					
Dn	222	CCACCACTTAAATAGACGACACAAATGATGATGCTGGATGAGCTCTTACAGATGACCA	281					
Qy	180	GGAGATGACCCCAAAAAACAAGAGAGGACAGAGAAGATCATCAAGAACTTATCAAGAAG	239					
Dn	282	GGAGATGACCCCAAAAAACAAGAGAGGACAGAGAAGATCATCAAGAACTTATCAAGAAG	341					
Qy	240	TCATCAAGCTGGCCATTCTTTATAGAAATATCATGTTTAAATCAAGATGACCTAGCATTTGA	299					
Dn	342	TCATCAAGCTGGCCATTCTTTATAGAAATATCATGTTTAAATCAAGATGACCTAGCATTTGA	401					
Qy	300	TGAGAAATTTAAAGAAAAGTTTCATACGCTTGTCTATGACCCGTGTACATTTCCATCAGG	359					
Dn	402	TGAGAAATTTAAAGAAAAGTTTCATACGCTTGTCTATGACCCGTGTACATTTCCATCAGG	461					
Qy	360	TGATTTATACCTTTGACCCGGAATGTGTTATCCAGGCTGTTAAATGAAATGACAGAGATGC	419					
Dn	462	TGATTTATACCTTTGACCCGGAATGTGTTATCCAGGCTGTTAAATGAAATGACAGAGATGC	521					
Qy	420	TGACCAAAATCATTTACAGCCGCACTCACTGCAAGTCACATGGAACGGGTTAATATGTCT	479					
Dn	522	TGACCAAAATCATTTACAGCCGCACTCACTGCAAGTCACATGGAACGGGTTAATATGTGT	581					
Qy	480	TTGATCATTTTTCAGATTTTGGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTGA	539					
Dn	582	TTGATCATTTTTCAGATTTTGGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTGA	641					
Qy	540	AACCCCACTTACAAAACCTATGTGATGGTATCAACAAAATGTTGATGAGAAGAACATAT	599					
Dn	642	AACCCCACTTACAAAACCTATGTGATGGTATCAACAAAATGTTGATGAGAAGAACATAT	701					
Qy	600	GA 601						
Dn	702	GA 703						
RESULT 5								
LOCUS	BI819214	813 bp	mRNA	linear	EST 04-OCT-2001			
DEFINITION	603034630F1 NIH_MGC_115	Homo sapiens	cDNA clone	IMAGE:5175702	5',			
ACCESSION	BI819214							
VERSION	BI819214.1	GI:15930764						
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.							
TITLE	1 (bases 1 to 813)							
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.							
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)							
	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: sgabds-remail.nih.gov							
	Tissue Procurement: Life Technologies, Inc.							
	cDNA Library Preparation: Life Technologies, Inc.							
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)							
	DNA Sequencing by: Incyte Genomics, Inc.							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LNL at:							
	http://image.lnl.gov							
	plate: L14M11437 row: 1 column: 07							
	High quality sequence stop: 809.							
FEATURES	Location/Qualifiers							
SOURCE	1. 813							
	/organism="Homo sapiens"							
	/mol_type="mRNA"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:5175702"							
	/lab_host="DH10B"							

Db 602 AACCCACTTACAAAATCTATGTATGATGATCAACAAATGTTGATGAAGACATAT 661
QY 600 GA 601
Db 662 GA 663

RESULT 7
BI752550 876 bp mRNA linear EST 25-SEP-2001
LOCUS 603021969F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5192561 5',
DEFINITION mRNA sequence.
ACCESSION BI752550
VERSION BI752550.1 GI:15744128
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 876)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@rs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
http://image.llnl.gov
Plate: LLM11481 row: j column: 18
High quality sequence stop: 813.
Location/Qualifiers
1..876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5192561"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_114"
/note="Organ: brain; Vector: PCMV-SPORE6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 95.5%; Score 574.2; DB 4; Length 876;
Best Local Similarity 99.2%; Pred. No. 3.5e-148;
Matches 558; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 CTTACGCTCCGCGCGCGCTGC-CACTCTCCGATGCCACAGATGCTTTAATTCA 59
Db 83 CTTACGCTCCGCGCGCGCGCTGC-CACTCTCCGATGCCACAGATGCTTTAATTCA 142

QY 60 AAAAAGCTGCGCTTCAGGCAAAAAGATCTTGGTAAATGCTGCCAATTCATG 119
Db 143 AAAAAGCTGCGCTTCAGGCAAAAAGATCTTGGTAAATGCTGCCAATTCATG 202

QY 120 CCAACCACTTAATAGACGACACAAGTAGTGCTGATGAGCTTACAGAGTACCA 179
Db 203 CCAACCACTTAATAGACGACACAAGTAGTGCTGATGAGCTTACAGAGTACCA 262

QY 180 GGGAGTACACCCAAAACAGAGAGGCGAGAGATCATCAAGAACTTCATCAAGCAG 239
Db 263 GGGAGTACACCCAAAACAGAGAGGCGAGAGATCATCAAGAACTTCATCAAGCAG 322

QY 240 TCATCAGCTGCGCATTCCTTTATAGAAATACGTTAATCAAGATGATGATGCA 299
Db 323 TCATCAGCTGCGCATTCCTTTATAGAAATACGTTAATCAAGATGATGATGCA 382

QY 300 TGGAGAAATTTAAGAAAGATTTCATGCTTGTATGACCGTGTGCTTCCATCAG 359
Db 383 TGGAGAAATTTAAGAAAGATTTCATGCTTGTATGACCGTGTGCTTCCATCAG 442

QY 360 TGGATTAATCTTTGACCGGAATGTGTTATCCAGGCTGTAAATGAAGAGATGC 419
Db 443 TGGATTAATCTTTGACCGGAATGTGTTATCCAGGCTGTAAATGAAGAGATGC 502

QY 420 TGCACCAATCATTTACAGCGCCACCTCAGTCCAAATCATGAGCGGTTAATGCT 479
Db 503 TGCACCAATCATTTACAGCGCCACCTCAGTCCAAATCATGAGCGGTTAATGCT 562

QY 480 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATCC-TTTTGGAAATTT 538
Db 563 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATCC-TTTTGGAAATTT 622

QY 539 AAACCCCACTTACAAAATCTATGTATGATGATCAACAAATGTTGATGAAGACATA 598
Db 623 AAACCCCACTTACAAAATCTATGTATGATGATCAACAAATGTTGATGAAGACATA 682

QY 599 TGA 601
Db 683 TGA 685

RESULT 8
CN305327 712 bp mRNA linear EST 16-MAY-2004
LOCUS CN305327
DEFINITION 17000532623591 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN305327
VERSION CN305327.1 GI:47321741
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 712)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guejler, K., Rao, M.S., Mandalam, R.,
Lebowke, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 712 Std Error: 0.00.
Location/Qualifiers
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cisse_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_1ib="GRN_ES"
/note="oligo dT primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Query Match 93.9%; Score 564.4; DB 7; Length 712;
Best Local Similarity 99.8%; Pred. No. 1.8e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGCGCGCTTACGACACAAAAGATCTTGG 95

```

Db      82  TGGCAGAGATGCTTTAAATTCAGAAACCTGGCCGCTTCAGGCAAGAAAGATCTTGG 141
Qy      96  GTTAAATGCTGTCCAAATCCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 155
Db      142  GTTAAATGCTGTCCAAATCCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 201
Qy      156  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAACCAAGAGAGAGAGAGA 215
Db      202  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAACCAAGAGAGAGAGAGA 261
Qy      216  TCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATTTCTTATAGAAATATCAGT 275
Db      262  TCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATTTCTTATAGAAATATCAGT 321
Qy      276  TTAATCAAGATGAGCTAGTATGATGAGAAATTTAAGAAAGATTCATCAGCTTCTA 335
Db      322  TTAATCAAGATGAGCTAGTATGATGAGAAATTTAAGAAAGATTCATCAGCTTCTA 381
Qy      336  TGACCGTGTCACTTTTCATCAGTGTATATACCTTTGACCGGAATGTGTATCCAGGC 395
Db      382  TGACCGTGTCACTTTTCATCAGTGTATATACCTTTGACCGGAATGTGTATCCAGGC 441
Qy      396  TGTAAATGATGAGAGAGATGCTGACCAATCATTCAGCGCACCTGACCTGCGCAAGT 455
Db      442  TGTAAATGATGAGAGAGATGCTGACCAATCATTCAGCGCACCTGACCTGCGCAAGT 501
Qy      456  CACATGAGCGGGTAAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 515
Db      502  CACATGAGCGGGTAAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 561
Qy      516  TGTATATCCTTTTGGAAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 575
Db      562  TGTATATCCTTTTGGAAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 621
Qy      576  AAATGTTGATGAAGAAACATATGA 601
Db      622  AAATGTTGATGAAGAAACATATGA 647

```

```

RESULT 9 718 bp  mRNA  linear  EST 25-MAR-2004
AL549492  LOCUS
DEFINITION  AL549492 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS001044YG13 5-PRIME, mRNA sequence.
ACCESSION  AL549492
VERSION    AL549492.3  GI:45749887
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 718)
REFERENCE  1  W.B., Gruber,C., Jesse,J. and Polyes,D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
COMMENT    On Feb 15, 2001 this sequence version replaced gi:31271310.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Creneau, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3485.f

```

```

FEATURES
source
1..718
/organism="Homo sapiens"
/mol_type="mRNA"
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS001044AD070P1c=3485.f.
Location/Qualifiers
1..718
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="CS001044YG13"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

Query Match 93.9%; Score 564.4; DB 1; Length 718;
Best Local Similarity 99.8%; Pred. No. 1,8e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      36  TGGCCAGAGTGTCTTTAAATTCAGAAACCTGGCCGCTTCAGGCAAGAAAGATCTTGG 95
Db      1  TGGCCAGAGTGTCTTTAAATTCAGAAACCTGGCCGCTTCAGGCAAGAAAGATCTTGG 60
Qy      96  GTTAAATGCTGTCCAAATCCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 155
Db      61  GTTAAATGCTGTCCAAATCCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 120
Qy      156  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAACCAAGAGAGAGAGAGA 215
Db      121  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAACCAAGAGAGAGAGAGA 180
Qy      216  TCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATTTCTTATAGAAATATCAGT 275
Db      181  TCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATTTCTTATAGAAATATCAGT 240
Qy      276  TTAATCAAGATGAGCTAGTATGATGAGAAATTTAAGAAAGATTCATCAGCTTCTA 335
Db      241  TTAATCAAGATGAGCTAGTATGATGAGAAATTTAAGAAAGATTCATCAGCTTCTA 300
Qy      336  TGACCGTGTCACTTTTCATCAGTGTATATACCTTTGACCGGAATGTGTATCCAGGC 395
Db      301  TGACCGTGTCACTTTTCATCAGTGTATATACCTTTGACCGGAATGTGTATCCAGGC 360
Qy      396  TGTAAATGATGAGAGAGATGCTGACCAATCATTCAGCGCACCTGACCTGCGCAAGT 455
Db      361  TGTAAATGATGAGAGAGATGCTGACCAATCATTCAGCGCACCTGACCTGCGCAAGT 420
Qy      456  CACATGAGCGGGTAAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 515
Db      421  CACATGAGCGGGTAAATATGCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 480
Qy      516  TGTATATCCTTTTGGAAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 575
Db      481  TGTATATCCTTTTGGAAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 540
Qy      576  AAATGTTGATGAAGAAACATATGA 601
Db      541  AAATGTTGATGAAGAAACATATGA 566

```

```

RESULT 10 731 bp  mRNA  linear  EST 06-JUN-2003
CD521721  LOCUS
DEFINITION  CD521721 NIH-MGC 14353034 NIH-MGC_191 Homo sapiens cDNA clone
            IMAGE:30413002 5', mRNA sequence.
ACCESSION  CD521721
VERSION    CD521721.1  GI:31453439
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 731)
REFERENCE  1  NIH-MGC http://mgi.cni.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Daniela S. Gerhard, ph.D.
            Office of Cancer Genomics

```

National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgsabbs-remail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: NDCM207 row: e column: 11
High quality sequence stop: 603.

FEATURES

source

```
1. 731
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30413002"
/tissue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 191"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggcccatatggcc);
Site 2: SfiI (ggcgccctcgcc); Library is oligo-dT primed
and directionally cloned. PBMC - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGGCCATTATGCCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGGCGGAGGCGCCAGCAGG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
```

ORIGIN

Query Match 93.9%; Score 564.4; DB 6; Length 731;
Best Local Similarity 99.8%; Pred. No. 1.8e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCACAAGAGATCTTGG 95
31 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCACAAGAGATCTTGG 90
96 GTAATATGATGCTTCAATCATGCGCACCCTTAATGAGCGACACAGATGAGAGTGC 155
91 GAAAAATGATGCTTCAATCATGCGCACCCTTAATGAGCGACACAGATGAGAGTGC 150
156 TGGATGAGCTTACAGAGTGAAGCAGGAGTACACCAAAACAAGAGAGGAGAGAGA 215
151 TGGATGAGCTTACAGAGTGAAGCAGGAGTACACCAAAACAAGAGAGGAGAGAGA 210
216 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCACT 275
211 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCACT 270
276 TTAATCAAGATGAGTGAAGTGAAGAAATTTAAGAAAGATTCATCACTGCTGA 335
271 TTAATCAAGATGAGTGAAGTGAAGAAATTTAAGAAAGATTCATCACTGCTGA 330
336 TGAACGTGCTCAATTCATCAAGTGAATATACCTTTGACCGGAATGTTATCCAGG 395
331 TGAACGTGCTCAATTCATCAAGTGAATATACCTTTGACCGGAATGTTATCCAGG 390
396 TGTAAATGATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCATGCGCAAGT 455
391 TGTAAATGATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCATGCGCAAGT 450
456 CACATGAGAGGGTAAATATGCTTGAATCAATTTTCAGATGAGAAATTTTGGCTGCC 515
451 CACATGAGAGGGTAAATATGCTTGAATCAATTTTCAGATGAGAAATTTTGGCTGCC 510
516 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATATCAACA 575
```

```
|||||
Db 511 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATATCAACA 570
Qy 576 AATGTTGATGAAGAGAAATATATGA 601
Db 571 AATGTTGATGAAGAGAAATATATGA 596
```

RESULT 11

```
AL527566
LOCUS AL527566 897 bp mRNA linear EST 24-MAR-2004
DEFINITION AL527566 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC024YA23 5-PRIME, mRNA sequence.
ACCESSION AL527566
VERSION AL527566.3 GI:45702664
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 897)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31065417.
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3485.F
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DC024YA12QPlc=3485.f.
```

FEATURES

source

```
1. 897
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC024YA23"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 93.9%; Score 564.4; DB 1; Length 897;
Best Local Similarity 99.8%; Pred. No. 1.9e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCACAAGAGATCTTGG 95
1 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCACAAGAGATCTTGG 60
96 GTAATATGATGCTTCAATCATGCGCACCCTTAATGAGCGACACAGATGAGAGTGC 155
91 GAAAAATGATGCTTCAATCATGCGCACCCTTAATGAGCGACACAGATGAGAGTGC 150
156 TGGATGAGCTTACAGAGTGAAGCAGGAGTACACCAAAACAAGAGAGGAGAGAGA 215
121 TGGATGAGCTTACAGAGTGAAGCAGGAGTACACCAAAACAAGAGAGGAGAGAGA 180
216 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCACT 275
181 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCACT 240
276 TTAATCAAGATGAGTGAAGAAATTTAAGAAAGATTCATCACTGCTGCTGA 335
```


Db 241 TTAATCAAGATGAGTAGCATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 300
 QY 336 TGACCGTGTGATGTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 395
 Db 301 TGACCGTGTGATGTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 360
 QY 396 TGTAAATGAATGAGAGATGCTGACCAAAATCAATTGACCGGCACTCAGTGCAGT 455
 Db 361 TGTAAATGAATGAGAGATGCTGACCAAAATCAATTGACCGGCACTCAGTGCAGT 420
 QY 456 CACATGAGCGGTTAAATATGCTTGTATCATTTTCAGATGTGAATTTTGGCTGCT 515
 Db 421 CACATGAGCGGTTAAATATGCTTGTATCATTTTCAGATGTGAATTTTGGCTGCT 480
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGTGATCAACA 575
 Db 481 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGTGATCAACA 540
 QY 576 AAATGTTGATGAAGAAACATATGA 601
 Db 541 AAATGTTGATGAAGAAACATATGA 566

RESULT 12
 BM919223 1055 bp mRNA linear EST 12-MAR-2002
 LOCUS BM919223
 DEFINITION AGENCOURT_6715690 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748442
 5', mRNA sequence.

ACCESSION BM919223
 VERSION BM919223.1 GI:19369602
 KEYWORDS EST
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Makiyoda; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Bacteria; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM12776 row: d column: 11
 High quality sequence stop: 677.
 Location/Qualifiers

FEATURES
 source 1..1055
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5748442"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_120"
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 93.8%; Score 564.4; DB 5; Length 1055;
 Best Local Similarity 99.8%; Pred. No. 2e-145;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTAAATTCGAAAACCTGGCGCTTACAGCACAAGAAAGATCTTG 95
 Db 33 TGGCCACAGATGCTTAAATTCGAAAACCTGGCGCTTACAGCACAAGAAAGATCTTG 92
 QY 96 GTAAATGATGTCATCAATCATGCGCACCCTTAATAGACACAAAGATGAGGTGC 155
 Db 93 GTAAATGATGTCATCAATCATGCGCACCCTTAATAGACACAAAGATGAGGTGC 152
 QY 156 TGGATGAGCTTACAGAGTACCGAGAGTACACCCAAAACAAGAGAGGACAGAGA 215
 Db 153 TGGATGAGCTTACAGAGTACCGAGAGTACACCCAAAACAAGAGAGGACAGAGA 212
 QY 216 TCAATCAAGAACTTCATCAAGACAGTATCAAGTGGCCATTCTTATAGAAATACAGT 275
 Db 213 TCAATCAAGAACTTCATCAAGACAGTATCAAGTGGCCATTCTTATAGAAATACAGT 272
 QY 276 TTAATCAAGATGAGTAGTATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCTA 335
 Db 273 TTAATCAAGATGAGTAGTATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCTA 332
 QY 336 TGACCGTGTGATGTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 395
 Db 333 TGACCGTGTGATGTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 392
 QY 396 TGTAAATGAATGAGAGATGCTGACCAAAATCAATTGACCGGCACTCAGTGCAGT 455
 Db 393 TGTAAATGAATGAGAGATGCTGACCAAAATCAATTGACCGGCACTCAGTGCAGT 452
 QY 456 CACATGAGCGGTTAAATATGCTTGTATCATTTTCAGATGTGAATTTTGGCTGCT 515
 Db 453 CACATGAGCGGTTAAATATGCTTGTATCATTTTCAGATGTGAATTTTGGCTGCT 512
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGTGATCAACA 575
 Db 513 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGTGATCAACA 572
 QY 576 AAATGTTGATGAAGAAACATATGA 601
 Db 573 AAATGTTGATGAAGAAACATATGA 598

RESULT 13
 CR611795 1610 bp mRNA linear HTC 21-JUL-2004
 LOCUS CR611795
 DEFINITION full-length cDNA clone CS01057YA07 of Placenta Cot 25-normalized
 of Homo sapiens (human).
 ACCESSION CR611795
 VERSION CR611795.1 GI:50492602
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Makiyoda; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Bacteria; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/
 COMMENT Full-length cDNA libraries and normalization
 Contact: Feng Liang Email: fliang@life.techn.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 1610)
 Genoscope.

REFERENCE
 AUTHORS Direct Submision
 TITLE Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES
 source 1..1610
 Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1057YA07"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 93.9%; Score 564.4; DB 3; Length 1610;
Best Local Similarity 99.8%; Pred. No. 2.2e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTTGAGGACAAAAGAGATCTTGG 95
DB 1 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTTGAGGACAAAAGAGATCTTGG 60
QY 96 GTTAAATGGTGTCCAAATCCATGCGCAACCTTAATAGACACACAGTAGTAGTGC 155
DB 61 GTTAAATGGTGTCCAAATCCATGCGCAACCTTAATAGACACACAGTAGTAGTGC 120
QY 156 TGGATGAGCTCTACAGAGTGACCGAGGGAGTACCCGAAAACAAAGAGAGGAGAGAGA 215
DB 121 TGGATGAGCTCTACAGAGTGACCGAGGGAGTACCCGAAAACAAAGAGAGGAGAGAGA 180
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTATAGGAATATCACT 275
DB 181 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTATAGGAATATCACT 240
QY 276 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 335
DB 241 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 300
QY 336 TGACCGTGTGCTTCCATCAGGTGATTAATACCTTTCGCGGAATGTTATCCAGGC 395
DB 301 TGACCGTGTGCTTCCATCAGGTGATTAATACCTTTCGCGGAATGTTATCCAGGC 360
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCATCTGCCAAGT 455
DB 361 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCATCTGCCAAGT 420
QY 456 CACATGAGAGGGTTAATAATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
DB 421 CACATGAGAGGGTTAATAATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 480
QY 516 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAACA 575
DB 481 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAACA 540
QY 576 AAATGTTGATGAAGAACAATATGA 601
DB 541 AAATGTTGATGAAGAACAATATGA 566

RESULT 14
CR615621 1733 bp mRNA linear HTC 21-JUL-2004
LOCUS CR615621
DEFINITION full-length cDNA clone CS0DC024YA23 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
ACCESSION CR615621
VERSION CR615621.1 GI:50496428
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1733)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT : Feng Liang Email : fliang@life.uchicago.edu
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1733)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

1..1733

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC024YA23"
/issue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 93.9%; Score 564.4; DB 3; Length 1733;
Best Local Similarity 99.8%; Pred. No. 2.2e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTTGAGGACAAAAGAGATCTTGG 95
DB 1 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTTGAGGACAAAAGAGATCTTGG 60
QY 96 GTTAAATGGTGTCCAAATCCATGCGCAACCTTAATAGACACACAGTAGTAGTGC 155
DB 61 GTTAAATGGTGTCCAAATCCATGCGCAACCTTAATAGACACACAGTAGTAGTGC 120
QY 156 TGGATGAGCTCTACAGAGTGACCGAGGGAGTACCCGAAAACAAAGAGAGGAGAGAGA 215
DB 121 TGGATGAGCTCTACAGAGTGACCGAGGGAGTACCCGAAAACAAAGAGAGGAGAGAGA 180
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTATAGGAATATCACT 275
DB 181 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTATAGGAATATCACT 240
QY 276 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 335
DB 241 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCTA 300
QY 336 TGACCGTGTGCTTCCATCAGGTGATTAATACCTTTCGCGGAATGTTATCCAGGC 395
DB 301 TGACCGTGTGCTTCCATCAGGTGATTAATACCTTTCGCGGAATGTTATCCAGGC 360
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCATCTGCCAAGT 455
DB 361 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCATCTGCCAAGT 420
QY 456 CACATGAGAGGGTTAATAATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
DB 421 CACATGAGAGGGTTAATAATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 480
QY 516 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAACA 575
DB 481 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGATCAACA 540
QY 576 AAATGTTGATGAAGAACAATATGA 601
DB 541 AAATGTTGATGAAGAACAATATGA 566

RESULT 15
CR615131 1754 bp mRNA linear HTC 21-JUL-2004
LOCUS CR615131
DEFINITION full-length cDNA clone CSOD1081Y011 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR615131
VERSION CR615131.1 GI:50495938
KEYWORDS HTC; CNSLT_cDNA.

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1754)
AUTHORS	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
REMARK	Contact : Feng Liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1754)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
JOURNAL	Web : www.genoscope.cns.fr)
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	Location/Qualifiers
source	1..1754
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS001081Y011"
	/tissue_type="Placenta
	/plasmid="pCMVSPORT_6"
ORIGIN	
Query Match	93.9%; Score 564.4; DB 3; Length 1754;
Best Local Similarity	99.8%; Pred. No. 2.2e-145;
Matches	565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	36 TGGCCACGATGCTCTTTAATTCGAAACCTGGCCGTTCAAGCACAAAGAATCTTGG 95
Db	1 TGGCCACGATGCTCTTTAATTCGAAACCTGGCCGTTCAAGCACAAAGAATCTTGG 60
QY	96 GTAAATAGTGATGTCGAAATCCATCGCCACACCTTAATAGAGACACAAAGTAGAGTGC 155
Db	61 GTAAATAGTGATGTCGAAATCCATCGCCACACCTTAATAGAGACACAAAGTAGAGTGC 120
QY	156 TGGATGAGCTTCAAGAGTGAACGAGGAGTACACCCAAACAGAGAGGACAGAGA 215
Db	121 TGGATGAGCTTCAAGAGTGAACGAGGAGTACACCCAAACAGAGAGGACAGAGA 180
QY	216 TCATCAGAACCTCATCAAGAGTCAATCAAGCTGGCCATCTTTATGAAATATCAGT 275
Db	181 TCATCAGAACCTCATCAAGAGTCAATCAAGCTGGCCATCTTTATGAAATATCAGT 240
QY	276 TTATTCAGAGTACGACTGACATTGATGGAAGAAATTAGAAAGAGTTCATCAGCTTGTA 335
Db	241 TTATTCAGAGTACGACTGACATTGATGGAAGAAATTAGAAAGAGTTCATCAGCTTGTA 300
QY	336 TGAACGATGATGATTTCCATCAGAGTGAATTAACCTTTGACCGGAATGTTATCCAGGC 395
Db	301 TGAACGATGATGATTTCCATCAGAGTGAATTAACCTTTGACCGGAATGTTATCCAGGC 360
QY	396 TGTTAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCACTGCCAAGT 455
Db	361 TGTTAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCACTGCCAAGT 420
QY	456 CACATGACCGGGTAAATTAATGCTTTGATTCATTTTTCAGATTGGAATTTTGGCTGCT 515
Db	421 CACATGACCGGGTAAATTAATGCTTTGATTCATTTTTCAGATTGGAATTTTGGCTGCT 480
QY	516 TGTATTAATCCTTTGGGAATTTTAAACCCCACTTACAAAACTAATGATGATTCACA 575
Db	481 TGTATTAATCCTTTGGGAATTTTAAACCCCACTTACAAAACTAATGATGATTCACA 540
QY	576 AAATGTTGGATGAAGACATATGCA 601

Db	541	AAATGTTGATGAAGACATATGA	566
RESULT 16	AL550457	885 bp	mrna
LOCUS	AL550457		linear
DEFINITION	AL550457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI057YA07 5-PRIME, mRNA sequence.		EST 25-MAR-2004
ACCESSION	AL550457		
VERSION	AL550457.3	GI:45750828	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 885)		
TITLE	Li, W.B., Gruber, C., Jessup, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	On Feb 15, 2001 this sequence version replaced gi:31272274.		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE		
	Email: seqrel@genoscope.cns.fr, web : www.genoscope.cns.fr		
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3485.f		
	For more information about this cluster, see		
	http://www.genoscope.cns.fr/cdna?cs=CSODI057YA04Op1&c=3485.f.		
FEATURES	location/Qualifiers		
SOURCE	1..885		
	/organism="Homo sapiens"		
	/mol_type="mrna"		
	/db_xref="taxon:9606"		
	/clone="CSODI057YA07"		
	/library="PLACENTA COT 25-NORMALIZED"		
	/issue_type="Homo sapiens PLACENTA COT 25-NORMALIZED"		
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
ORIGIN			
	Query Match 93.6%; Score 562.8; DB 1; Length 885;		
	Best Local Similarity 99.6%; Pred. No. 5.2e-145;		
	Matches 564; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	36	TGGCCACAGATGCTTTAATTCGAAAACTGGCGCTTCAGGCACAAAGAAGTCTTGG	95
Db	1	TGGCCACAGATGCTTTAATTCGAAAACTGGCGCTTCAGGCACAAAGAAGTCTTGG	60
Qy	96	GTAATAATGATGCTCAATCATCGCCACCACTTAATAGACACACAAGTAGTAGGTGC	155
Db	61	GTAATAATGATGCTCAATCATCGCCACCACTTAATAGACACACAAGTAGTAGGTGC	120
Qy	156	TGATGAGCTCTACAGAGTACACGAGGAGTACACCCAAAACAAGAGGAGCAGAGA	215
Db	121	TGATGAGCTCTACAGAGTACACGAGGAGTACACCCAAAACAAGAGGAGCAGAGA	180
Qy	216	TCAACAAGAACCTCATTAAGACAGTATCAAGCTGGCCATCTTTATAGAAATATCA	275
Db	181	TCAACAAGAACCTCATTAAGACAGTATCAAGCTGGCCATCTTTATAGAAATATCA	240
Qy	276	TTAATCAAGATGAGCTAGCTATGAGAGAAATTTAAGAGAAATTCATCAGCTTGCTA	335
Db	241	TTAATCAAGATGAGCTAGCTATGAGAGAAATTTAAGAGAAATTCATCAGCTTGCTA	300
Qy	336	TGACCGTGTGCACTTTCATCAGGTGATTAATCTTGAACCGGAATGTGTATCCAGGC	395

Db 301. TGACCGTGTGCACTTTCATCACTGATGATTAATCACTTTGACCGGAAATGTTATCCAGGC 360
Qy 396 TGTAAATGATGACAGAGATGCTGACCAATCACTTACAGCGCACTCACTGCAAGT 455
Db 361 TGTAAATGATGACAGAGATGCTGACCAATCACTTACAGCGCACTCACTGCAAGT 420
Qy 456 CACATGACGGGTAAATGATGCTTTCATCACTTTTCAGATGTAATTTTGGCTGCT 515
Db 421 CACATGACGGGTAAATGATGCTTTCATCACTTTTCAGATGTAATTTTGGCTGCT 480
Qy 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 575
Db 481 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 540
Qy 576 AATGTTGGATGAAGAACATATGA 601
Db 541 AATGTTGGATGAAGAACATATGA 566

RESULT 17
AL554168 1045 bp mRNA linear EST 30-MAR-2004
LOCUS AL554168 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1081Y011 5-PRIME, mRNA sequence.
ACCESSION AL554168
VERSION AL554168.3 GI:45858929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31275981.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3485.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1081Y011&c=3485.f.

FEATURES
source location/Qualifiers
1..1045
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081Y011"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/tissue_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 93.6%; Score 562.4; DB 1; Length 1045;
Best Local Similarity 99.5%; Pred. No. 7e-145; Indels 0; Gaps 0;

Matches 563; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 36 TGGCCACAGATGCTTTAATTCAAAAACCTGGCCGCTTCAGGACAAAAAGAGATCTTG 95
Db 1 TGGCCACAGATGCTTTAATTCAAAAACCTGGCCGCTTCAGGACAAAAAGAGATCTTG 60
Qy 96 GTAAAAATGATGCTTCAATTCATGCGACCACTTAATAGACGACAAAGTAGAGAGTGC 155
Db 61 GTAAAAATGATGCTTCAATTCATGCGACCACTTAATAGACGACAAAGTAGAGAGTGC 120

Qy 156 TGTATAGCTCTACAGAGTACAGGAGTACACCCAAAAACAAGAGAGGACAGAGA 215
Db 121 TGTATAGCTCTACAGAGTACAGGAGTACACCCAAAAACAAGAGAGGAGAGA 180
Qy 216 TCATCAAGAACCTCATCAAGACATCATCAAGCTGCGCACTTTTATAGAAATCACT 275
Db 181 TCATCAAGAACCTCATCAAGACATCATCAAGCTGCGCACTTTTATAGAAATCACT 240
Qy 276 TTAATCAAGATGAGCTAGATGATGAGAAATTTAAGAAAGAAATGATGAGTGC 335
Db 241 TTAATCAAGATGAGCTAGATGATGAGAAATTTAAGAAAGAAATGATGAGTGC 300
Qy 336 TGACCGTGTGCACTTTCATCACTGATGATTAATCACTTTGACCGGATGTTATCAGGC 395
Db 301 TGACCGTGTGCACTTTCATCACTGATGATTAATCACTTTGACCGGATGTTATCAGGC 360
Qy 396 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGGCGCACTCACTGCAAGT 455
Db 361 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGGCGCACTCACTGCAAGT 420
Qy 456 CACATGACGGGTAAATGATGCTTTCATCACTTTTCAGATGTAATTTTGGCTGCT 515
Db 421 CACATGACGGGTAAATGATGCTTTCATCACTTTTCAGATGTAATTTTGGCTGCT 480
Qy 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 575
Db 481 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 540
Qy 576 AATGTTGGATGAAGAACATATGA 601
Db 541 AATGTTGGATGAAGAACATATGA 566

RESULT 18
CB991862 785 bp mRNA linear EST 01-MAY-2003
LOCUS CB991862
DEFINITION AGENCOURT 13618130 NIH MGC 148 Homo sapiens cDNA clone
IMAGE:30337918 5', mRNA sequence.
ACCESSION CB991862
VERSION CB991862.1 GI:30286382
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hanson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNLT at:
http://image.llnl.gov
Plate: NDM363 row: 1 column: 23
High quality sequence stop: 629.

FEATURES
source location/Qualifiers
1..785
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30337918"
/tissue_type="pre-ecampic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 148"
/note="Organ: Placenta; Vector: pBluescriptR; Site: 1:
all-XhoI; Site_2: BamH; Library is oligo-dT primed and

LOCUS CR548728 779 bp mRNA linear EST 12-JUL-2004
 DEFINITION DKFZp469E1232_r1_469 (synonym: pkid1) Pongo pygmaeus cDNA clone
 ACCESSION DKFZp469E1232_5', mRNA sequence.
 VERSION CR548728
 KEYWORDS EST.
 SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 1 (bases 1 to 779)
 Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amlid, C., Osanger, A., Fobo, G., Han, M., and Wiemann, S.
 Pongo pygmaeus mRNA (Bahr, A., Lauber, J., Mewes, H.W., et al.)
 JOURNAL Unpublished (2004)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469E1232) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.
 Location/Qualifiers
 1..779
 /organism="Pongo pygmaeus"
 /mol_type="mRNA"
 /db_xref="taxon:9600"
 /clone="DKFZp469E1232"
 /tissue_type="kidney"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="469 (synonym: pkid1)"
 /note="Vector: pSPORT1_Sfi; Site_1: SfiI; Site_2: SfiIb"
 ORIGIN
 Query Match 92.6%; Score 556.4; DB 7; Length 779;
 Best Local Similarity 98.9%; Pred. No. 3e-143;
 Matches 560; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 36 TGGCCACAGATGCTTTAATTCAGAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTG 95
 DB 185 TGGCCACAGATGCTTTAATTCAGAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTG 244
 QY 96 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGGTG 155
 DB 245 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGGTG 304
 QY 156 TGGATGAGCTTACAGAGTGACCAAGGAGTACCCAAAACAAAGAGGAGAGAGA 215
 DB 305 TGGATGAGCTTACAGAGTGACCAAGGAGTACCTAAAACAAAGAGGAGAGAGA 364
 QY 216 TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGAAATTAATCAGT 275
 DB 365 TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGAAATTAATCAGT 424
 QY 276 TTAATCAAGTAGAGTAGAGTATGAGAGAAATTTAAGAAAGAGTTCAACAGCTTGCTA 335
 DB 425 TTAATCAAGTAGAGTAGAGTATGAGAGAAATTTAAGAAAGAGTTCAACAGCTTGCTA 484
 QY 336 TGAACGTGTGCTTTCATCAGAGTGAATTATACCTTTCACCGAAATGTGTTACAGGC 395
 DB 485 TGAACGTGTGCTTTCATCAGAGTGAATTATACCTTTCACCGAAATGTGTTACAGGC 544
 QY 396 TGTAAATGATGACAGAGATGCTGACCAATCAATTCATTCAGGCGCACTTCAGCAAGT 455
 DB 545 TGTAAATGATGACAGAGATGCTGACCAATCAATTCATTCAGGCGCACTTCAGCAAGT 604

QY 456 CACATGACGGGTATATATGCTTGCATCATTTTTCAGATGATGAAATTTTGCGTGCCT 515
 DB 605 CACATGACGGGTATATATGCTTGCATCATTTTTCAGATGATGAAATTTTGCGTGCCT 664
 QY 516 TGTATATCTCTTTGGGAAATTTTAAACCCCACTTACAAAACCTATGATGATGATACAA 575
 DB 665 TGTATATCTCTTTGGGAAATTTTAAACCCCACTTACAAAACCTATGATGATGATACAA 724
 QY 576 AATGTGTGATGACAGAAACATATGA 601
 DB 725 AATGTGTGATGACAGAAACATATGA 750
 RESULT 21
 LOCUS AL558873 559 bp mRNA linear EST 02-APR-2004
 DEFINITION AL558873 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 ACCESSION AL558873
 VERSION AL558873
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 559)
 Li, W.B., Gruber, C., Jessee, J., and Polyaes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:31283006.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3485.f
 For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?c=CS0DJ015DH02QPlc=3485.f>.
 Location/Qualifiers
 1..559
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ015YH04"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 Query Match 92.3%; Score 554.6; DB 1; Length 559;
 Best Local Similarity 98.6%; Pred. No. 8.8e-143;
 Matches 551; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
 QY 36 TGGCCACAGATGCTTTAATTCAGAAAACCTGGCCGTTTACAGGACAAAAGAGATCTTG 95
 DB 1 TGGCCACAGATGCTTTAATTCAGAAAACCTGGCCGTTTACAGGACAAAAGAGATCTTG 60
 QY 96 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGGTG 155
 DB 61 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGGTG 120
 QY 156 TGGATGAGCTTACAGAGTGACCAAGGAGTACCCAAAACAAAGAGGAGAGAGA 215
 DB 121 TGGATGAGCTTACAGAGTGACCAAGGAGTACCCAAAACAAAGAGGAGAGAGA 180

QY 216 TCATCAAGAACCTTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGGAATTAATAGT 275
 DB 181 TCATCAAGAACCTTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGGAATTAATAGT 240
 QY 276 TTAATCAAGATGAGCTAGCTATGATGAGAAATTTAAGAAAGATTTCATCAGCTGCTA 335
 DB 241 TTAATCAAGATGAGCTAGCTATGATGAGAAATTTAAGAAAGATTTCATCAGCTGCTA 300
 QY 336 TGACCGTGTGATGATTCATCAGGTGATTAACCTTTGACCGGAATGTGTTATCCAGGC 395
 DB 301 TGACCGTGTGATGATTCATCAGGTGATTAACCTTTGACCGGAATGTGTTATCCAGGC 360
 QY 396 TGTAAATGATGAGAGAGATGCTGACCAAAATCATTCAGCCGACCTCACTGCCAAGT 455
 DB 361 TGTAAATGATGAGAGAGATGCTGACCAAAATCATTCAGCCGACCTCACTGCCAAGT 420
 QY 456 CACATGAGACGGGTTAATATGTCCTTGTATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
 DB 421 CACATGAGACGGGTTAATATGTCCTTGTATCATTTTTCAGATTGTGAATTTTGGCTGCT 480
 QY 516 TGTATAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 575
 DB 481 TGTATAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 540
 QY 576 AAATGTTGATGAGAGAA 594
 DB 541 AAATGTTGATGAGAGAA 559

RESULT 22
 LOCUS B1759027 893 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603042814F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183355 5',
 mRNA sequence.
 ACCESSION B1759027
 VERSION B1759027.1 GI:15750605
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgsbds-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1457 row: k column: 04
 High quality sequence start: 5
 High quality sequence stop: 869.
 Location/Qualifiers

FEATURES

source
 1..893
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5183355"
 /lab_host="DH10B"
 /clone_lib="NIH MGC_116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb.

ORIGIN

Query Match 91.9%; Score 552.4; DB 4; Length 893;
 Best Local Similarity 99.6%; Pred. No. 4e-142;
 Matches 564; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."

QY 36 TGGCCACAGATGTCCTTAATTCAAAACCTGGCCCTTTCAGGCACAAAAGATCTTGG 95
 DB 95 TGGCCACAGATGTCCTTAATTCAAAACCTGGCCCTTTCAGGCACAAAAGATCTTGG 154
 QY 96 GTAAATGTCCTCAATCCATGCGCACCTTAATAGACACACAGTAGTAGGTGC 155
 DB 155 GTAAATGTCCTCAATCCATGCGCACCTTAATAGACACACAGTAGTAGGTGC 214
 QY 156 TGGATGAGCTCTACAGAGTACCAAGGAGTACCCAAAACAAGAGAGGACAGAGA 215
 DB 215 TGGATGAGCTCTACAGAGTACCAAGGAGTACCCAAAACAAGAGAGGACAGAGA 274
 QY 216 TCATCAAGAACCTTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGGAATTAATCAGT 275
 DB 275 TCATCAAGAACCTTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGGAATTAATCAGT 334
 QY 276 TTAATCAAGATGAGCTAGCTAGCTAGTATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGTCTA 335
 DB 335 TTAATCAAGATGAGCTAGCTAGCTAGTATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGTCTA 334
 QY 336 TGACCGTGTGATGATTCATCAGGTGATTAACCTTTGACCGGAATGTGTTATCCAGGC 395
 DB 395 TGACCGTGTGATGATTCATCAGGTGATTAACCTTTGACCGGAATGTGTTATCCAGGC 454
 QY 396 TGTAAATGATGAGAGAGATGCTGACCAAAATCATTCAGCGCACCTCACTGCCAAGT 455
 DB 455 TGTAAATGATGAGAGAGATGCTGACCAAAATCATTCAGCGCACCTCACTGCCAAGT 514
 QY 456 CACATGAGACGGGTTAATATGTCCTTGTATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
 DB 515 CACATGAGACGGGTTAATATGTCCTTGTATCATTTTTCAGATTGTGAATTTTGGCTGCT 573
 QY 516 TGTATAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 575
 DB 574 TGTATAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAACA 633
 QY 576 AAATGTTGATGAGAGAACTATGTA 601
 DB 634 AAATGTTGATGAGAGAACTATGTA 659

RESULT 23
 LOCUS CN648259 998 bp mRNA linear EST 13-MAY-2004
 DEFINITION IL10L10GEN_MCO_30108 Katze_MPB Macaca mulatta cDNA clone IBIDW:6795
 5' similar to Bases 72 to 998 highly similar to human TNFAIP8
 (Hs.17839), mRNA sequence.

ACCESSION CN648259
 VERSION CN648259.1 GI:47161702
 KEYWORDS EST.
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecoidea; Macaca.

REFERENCE
 AUTHORS Katze, M.G., Thomas, M., Korth, M., Tadonato, S.P. and Magness, C.L.
 TITLE Large-scale Rhesus Macaque cDNA Sequencing
 JOURNAL Unpublished (2003)
 COMMENT Contact: C. Magness
 Illumigen Biosciences Inc.
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400

Fax: 2063780408
 Email: cmagnes@illumigen.com
 Sequenced on 2004.03.27. 798 Q20 bases.
 PCR Primers
 FORWARD: CCTCTACTAAGGGAACAA
 BACKWARD: CACTATAGCGCAATTGGCTA
 Insert Length: 998 Std Error: 0.00
 Plate: CL000256 row: D column: 03
 Seq primer: CCTCTACTAAGGGAACAA
 POLYA=No.

FEATURES

source

Location/Qualifiers
 1. 998
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /clone="IBIDW:6795"
 /sex="male"
 /tissue_type="blood"
 /cell_type="PBMC"
 /dev_stage="adult"
 /lab_host="E. coli: SOLR"
 /clone_lib="Katzel MMPB"
 /note="Vector: Uni-ZAP XR, Site 1: EcoR I, Site 2: Xho I;
 Created from Stratagene ZAP-CDNA Synthesis kit (catalog
 #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit
 (Catalog #200450)"

ORIGIN

Query Match 91.2%; Score 548.4; DB 7; Length 998;
 Best Local Similarity 98.1%; Pred. No. 5.4e-141;
 Matches 555; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTGACGACAAAAAGAAATCTTGG 95
 |||||
 94 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTGACGACAAAAAGAAATCTTGG 153
 |||||
 96 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGACAGAGTGAAGTGC 155
 |||||
 154 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGACAGAGTGAAGTGC 213
 |||||
 156 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAAAAGAGAGGACAGAGA 215
 |||||
 214 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAAAAGAGAGGACAGAGA 273
 |||||
 216 TCATCAAGAACCTCAACAGAGTCAATCAAGTGGCCATTCTTAATAGAAATATCA 275
 |||||
 274 TCATCAAGAACCTCAACAGAGTCAATCAAGTGGCCATTCTTAATAGAAATATCA 333
 |||||
 276 TTAATCAAGATGAGTACAGATGATGAGAAATTTAAGAAAGATTCAGCTTGCT 335
 |||||
 334 TTAATCAAGATGAGTACAGATGATGAGAAATTTAAGAAAGATTCAGCTTGCT 393
 |||||
 336 TGAACGTGCTCACTTCCATCAAGTGAATTATACCTTTGACCGAAATGTGTATCC 395
 |||||
 394 TGAACGTGCTCACTTCCATCAAGTGAATTATACCTTTGACCGAAATGTGTATCC 453
 |||||
 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGGCCACTTCAGTCCA 455
 |||||
 454 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGGCCACTTCAGTCCA 513
 |||||
 456 CACATGAGCGGTTAATTAATGCTTGAATCATTTTCAAGTTGTAATTTGGTGGCT 515
 |||||
 514 CACATGAGCGGTTAATTAATGCTTGAATCATTTTCAAGTTGTAATTTGGTGGCT 573
 |||||
 516 TGTATATCTTTTGGAAATTTAAACCCCACTTACAAAAATATGATGATATCA 575
 |||||
 574 TGTATATCTTTTGGAAATTTAAACCCCACTTACAAAAATATGATGATATCA 633
 |||||
 576 AAATGTTGATGAAGAACATATGA 601
 |||||
 634 AAATGTTGATGAAGAACATATGA 659
 |||||

RESULT 24

AU135377

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AU135377 816 bp mRNA linear EST 02-ANG-2002
 AU135377 PLACE1 Homo sapiens CDNA clone PLACE1001920 5', mRNA
 sequence.
 AU135377
 AU135377.1 GI:10995916
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 816)
 Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saio, K., Kawai, Y.,
 Yamamoto, D., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
 Isogai, T.
 HRI human CDNA project
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@ri.co.jp
 HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix
 Research Institute; CDNA library construction; Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 Location/Qualifiers
 1. 816
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1001920"
 /tissue_type="Placenta"
 /clone_lib="PLACE1"
 /note="Vector: pME18SFL3"

FEATURES

source

Query Match 91.1%; Score 547.6; DB 1; Length 816;
 Best Local Similarity 98.9%; Pred. No. 8.5e-141;
 Matches 562; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

36 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTGACGACAAAAAGAAATCTTGG 95
 |||||
 102 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTGACGACAAAAAGAAATCTTGG 161
 |||||
 96 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGACAGAGTGAAGTGC 155
 |||||
 162 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGACAGAGTGAAGTGC 221
 |||||
 156 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAAAAGAGAGGACAGAGA 215
 |||||
 222 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAAAAGAGAGGACAGAGA 281
 |||||
 216 TCATCAAGAACCTCAACAGAGTCAATCAAGTGGCCATTCTTAATAGAAATATCA 275
 |||||
 282 TCATCAAGAACCTCAACAGAGTCAATCAAGTGGCCATTCTTAATAGAAATATCA 341
 |||||
 276 TTAATCAAGATGAGTACAGATGATGAGAAATTTAAGAAAGATTCAGCTTGCT 335
 |||||
 342 TTAATCAAGATGAGTACAGATGATGAGAAATTTAAGAAAGATTCAGCTTGCT 401
 |||||
 336 TGAACGTGCTCACTTCCATCAAGTGAATTATACCTTTGACCGAAATGTGTATCC 395
 |||||
 402 TGAACGTGCTCACTTCCATCAAGTGAATTATACCTTTGACCGAAATGTGTATCC 461
 |||||
 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGGCCACTTCAGTCCA 455
 |||||
 462 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGGCCACTTCAGTCCA 521
 |||||

ORIGIN

QY 456 CACATGACGGGTTAATATGCTTGTGATCATTTTTCAGATTGGAATTTTGGCTCCCT 515
 DB 522 CACATGACGGGTTAATATGCTTGTGATCATTTTTCAGATTGGAATTTTGGCTCCCT 581
 QY 516 TGTATTAATCCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGCTGTACAA 575
 DB 582 TGTATTAATCCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGCTGTACAA 641
 QY 576 AAT-CTTGATGAGAGAGAACTATGA 601
 DB 642 AATTTGTTGTGATGAGAGAACTATGA 669

RESULT 25
 BP302623 580 bp mRNA linear EST 17-SEP-2004
 LOCUS BP302623 Sugano cDNA library, macrophage Homo sapiens cDNA clone
 DEFINITION BP302623 mRNA sequence.
 ACCESSION BP302623 GI:52231583
 VERSION BP302623.1 GI:52231583
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 580)
 Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
 source
 1..580
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MP06324"
 /cell_type="macrophage"
 /clone_lib="Sugano cDNA library, macrophage"

ORIGIN
 Query Match 90.7% Score 545.4; DB 5; Length 580;
 Best Local Similarity 99.8%; Pred. No. 3.2e-140;
 Matches 546; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 55 TTCCAAAACCTGGCCGCTTCAGGACAAAGAGATCTGGTAAATGGTGTCAATC 114
 DB 1 TTCCAAAACCTGGCCGCTTCAGGACAAAGAGATCTGGTAAATGGTGTCAATC 60
 QY 115 CATGCCACCACTTAAATAGACACACAGTAGTAGCTGTGATGAGCTTTCAGAGT 174
 DB 61 CATGCCACCACTTAAATAGACACACAGTAGTAGCTGTGATGAGCTTTCAGAGT 120
 QY 175 GACACGGAGATGACCCAAAACAAAGAGAGGACAGAAATCATCAAGAACTCATCA 234
 DB 121 GACACGGAGATGACCCAAAACAAAGAGAGGACAGAAATCATCAAGAACTCATCA 180
 QY 235 GACAGTCATCAAGCTGGCCATCTTTTATAGAAATATCACTTAAATCAAGATGAGTAC 294
 DB 181 GACAGTCATCAAGCTGGCCATCTTTTATAGAAATATCACTTAAATCAAGATGAGTAC 240
 QY 295 ATTGATGAGAAATTTTAAAGAAAGAAAGTTCATCAGCTGTGATGAGCTGTTCAC 354
 DB 241 ATTGATGAGAAATTTTAAAGAAAGAAAGTTCATCAGCTGTGATGAGCTGTTCAC 300
 QY 355 TCAGGTGATTAATCTTTGACCGGAATGTGTTATCCAGGCTGTTAAATAGACAGAG 414
 DB 301 TCAGGTGATTAATCTTTGACCGGAATGTGTTATCCAGGCTGTTAAATAGACAGAG 360

QY 415 GATGCTGACCAAAATCATTTAGCGCCACCTCACTGCCAAATCAATGACAGCGGTTAATA 474
 DB 361 GATGCTGACCAAAATCATTTAGCGCCACCTCACTGCCAAATCAATGACAGCGGTTAATA 420
 QY 475 TGTCTTGAATCATTTTTCAGATTGTGAATTTTGGCGCTGTGATTAATCTTTGGGAA 534
 DB 421 TGTCTTGAATCATTTTTCAGATTGTGAATTTTGGCGCTGTGATTAATCTTTGGGAA 480
 QY 535 TTTTAAACCCCACTTACAAAACCTATGATGATGATCAACAAATGTTGATGAGAGAA 594
 DB 481 TTTTAAACCCCACTTACAAAACCTATGATGATGATCAACAAATGTTGATGAGAGAA 540
 QY 595 CATATGA 601
 DB 541 CATATGA 547

RESULT 26
 CA843219 596 bp mRNA linear EST 13-DEC-2002
 LOCUS CA843219
 DEFINITION CA843219.1 HR85.1a1et Homo sapiens cDNA clone IMAGE:6548934 5',
 similar to TR:095379 095379 MDC-3.13 ISOFORM 2. [2] TR:09UP47 ,
 mRNA sequence.
 ACCESSION CA843219 GI:26673548
 VERSION CA843219.1 GI:26673548
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 596)
 Melton,D., Brown,J., Kenty,G., Pernutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Maria,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
 Schmitt,A., Theising,B., Riter,E., Ronko,I., Bennett,T.,
 Cardenas,M., Gibbons,M., McCann,R., Cole,K., Tsagarisvili,R.,
 Williams,T., Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@hmp.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 461.

FEATURES
 source
 1..596
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6548934"
 /issue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85.1a1et"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
 NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permut Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 90.2%; Score 542.4; DB 6; Length 596;
 Best Local Similarity 99.6%; Pred. No. 2, 2e-139;
 Matches 554; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 CTTGACGCTCCGCGCGCGCTGCG-CGACTCTCCGATGGCCACAGATGCTTTAATTCGA 59
    |||||
DB 40 CTTCAACGCTCCGCGCGCGCGCTGCGCGCACTCTCCGATGGCCACAGATGCTTTAATTCGA 99
    |||||
QY 60 AAAACCTGCGCGCTTCGCGCAAAAAGAATCTTGGGTAAATNGTGTCCAAATCCATCG 119
    |||||
DB 100 AAAACCTGCGCGCTTCGCGCAAAAAGAATCTTGGGTAAATNGTGTCCAAATCCATCG 159
    |||||
QY 120 CCACCACTTTAATAGACGACACAAGTAGTAGGTGCTGATGAGCTCTACAGAGTAGCA 179
    |||||
DB 160 CCACCACTTTAATAGACGACACAAGTAGTAGGTGCTGATGAGCTCTACAGAGTAGCA 219
    |||||
QY 180 GGGAGTAGACCCCAAAACAAGAGGACAGAAAGATCATCAAGAACTTCATCAAGACAG 239
    |||||
DB 220 GGGAGTAGACCCCAAAACAAGAGGACAGAAAGATCATCAAGAACTTCATCAAGACAG 279
    |||||
QY 240 TCATCAAGCTGCGCATCTTTAATAGAAATATCAAGTTAATCAAGTAGCTAGCTATGA 289
    |||||
DB 280 TCATCAAGCTGCGCATCTTTAATAGAAATATCAAGTTAATCAAGTAGCTAGCTATGA 339
    |||||
QY 300 TGGAGAAATTTAAGAGAAAGTTCATCAAGCTGCTATGACCGTGTCAAGTTCCATCAG 359
    |||||
DB 340 TGGAGAAATTTAAGAGAAAGTTCATCAAGCTGCTATGACCGTGTCAAGTTCCATCAG 399
    |||||
QY 360 TGGATTAATCCTTTGACCGGAATGTGTATTCAGGCTGTTAAATGAATGACAGAGATGC 419
    |||||
DB 400 TGGATTAATCCTTTGACCGGAATGTGTATTCAGGCTGTTAAATGAATGACAGAGATGC 459
    |||||
QY 420 TGCACCAAAATCATTCAGCGCCACCTCATCGCCAAAGTCAATGACGGGTTAAATAGTCT 479
    |||||
DB 460 TGCACCAAAATCATTCAGCGCCACCTCATCGCCAAAGTCAATGACGGGTTAAATAGTCT 519
    |||||
QY 480 TTGATCATTTTTCAGATGTGATTTTGGCTGCTTGTATATCTTTTGGGATTTTA 539
    |||||
DB 520 TTGATCATTTTTCAGATGTGATTTTGGCTGCTTGTATATCTTTTGGGATTTTA 579
    |||||
QY 540 AACCCCACTTACAAA 555
    |||||
DB 580 AACCCCACTTACAAA 595
    |||||

```

RESULT 27
 BX405197 834 bp mRNA linear EST 03-MAY-2004
 LOCUS BX405197 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
 DEFINITION clone CS0DH001YB17 5-PRIME, mRNA sequence.
 ACCESSION BX405197
 VERSION BX405197.2 GI:46952678
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 834)
 Li,W.B., Gruber,C., Jesse,J. and Polyes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 15, 2003 this sequence version replaced gi:30762567.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 3485.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS0AH001CA09QPlc=3485.f>.

FEATURES

source

1..834

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH001YB17"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_11b="Homo sapiens T CELLS (JURKAT CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and EcoRV
sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

ORIGIN

Query Match 89.6%; Score 538.6; DB 5; Length 834;
 Best Local Similarity 97.4%; Pred. No. 2, 7e-138;
 Matches 552; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

```

QY 36 TGGCCACAGATGCTTTAATTCGAAACCTGCGCGCTTCAGGCGACAAAAAGAGATCTGG 95
    |||||
DB 1 TGGCCACAGATGCTTTAATTCGAAACCTGCGCGCTTCAGGCGACAAAAAGAGATCTGG 60
    |||||
QY 96 GTAAATGGTGTCCAAATTCATGCGCACACCTTAATGAGACACAAGTAGTAGTGC 155
    |||||
DB 61 GTAAATGGTGTCCAAATTCATGCGSACCACCTTAATGAGACACAAGTAGTAGTGC 120
    |||||
QY 156 TGGATGAGCTCTACAGAGTACAGGAGTACAGGAGTACACCAAAACAAGAGGACAGAGA 215
    |||||
DB 121 TGGATGAGCTCTACAGAGTACAGGAGTACAGGAGTACACCAAAACAAGAGGACAGAGA 180
    |||||
QY 216 TCATCAAGAACTCATCAAGACAGTATCAAGCTGCGCATTTTATAGAAATATCACT 275
    |||||
DB 181 TCATCAAGAACTCATCAAGACAGTATCAAGCTGCGCATTTTATAGAAATATCACT 240
    |||||
QY 276 TTAATCAAGATGAGCTAGATGATGAGAAATTTAAGAAAGTCAATCAAGCTTGCTGA 335
    |||||
DB 241 TTAATCAAGATGAGCTAGATGATGAGAAATTTAAGAAAGTCAATCAAGCTTGCTGA 300
    |||||
QY 336 TGAACGTGTGTCAGTTCCATCAGTGTGATTAAC-TTGAACCGGAATGTGTATTCAG 394
    |||||
DB 301 TGAACGTGTGTCAGTTCCATCAGTGTGATTAACCGGAATGTGTATTCAG 360
    |||||
QY 395 CTGTTAATGAATGACAGAGATGCTGCACCAATATCATTACGCGCACCTCACTGCCAAG 454
    |||||
DB 361 CTGTTAATGAATGACAGAGATGCTGCACCAATATCATTACGCGCACCTCACTGCCAAG 420
    |||||
QY 455 TCACATGACGCGGTTAATTAATGCTTTGATCAATTTTTCAGATTTGGAATTTTGGCTGCC 514
    |||||
DB 421 TCACATGACGCGGTTAATTAATGCTTTGATCAATTTTTCAGATTTGGAATTTTGGCTGCC 480
    |||||
QY 515 TTGTAATATCCTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAAC 574
    |||||
DB 481 TTGTAATATCCTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAAC 540
    |||||
QY 575 AAAATGTTGATGAAGAGAACTATGA 601
    |||||
DB 541 AAAATGTTGATGAAGAGAACTATGA 567
    |||||

```

RESULT 28

CD520304

747 bp mRNA linear EST 06-JUN-2003

LOCUS CD520304

AGENCOURT 14360208 NIH MGC 191 Homo sapiens cDNA clone

DEFINITION IMAGE:30410510 5', mRNA sequence.

ACCESSION CD520304

CD520304.1 GI:31452022

VERSION CD520304

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE 1 (bases 1 to 747)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: csabds-remail.nih.gov
 Tissue Procurement: Narayan Bhat
 CDNA Library Preparation: Clontech Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: NDCM200 row: m column: 15
 High quality sequence stop: 499.
 Location/Qualifiers

FEATURES

source

1. 747
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30410510"
 /tissue_type="Pooled"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_id="NIH MGC 191"
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcctggcc); Site 2: SfiI (ggccctcggcc); Library is oligo-dT primed and directionally cloned. PBMC - Peripheral Blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA adn Ionomycin. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGSCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGACGCGCCGACATG-3' (30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 89.6%; Score 538.4; DB 6; Length 747;
 Best Local Similarity 97.2%; Pred. No. 3e-138;
 Matches 548; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

36 TGCCCAAGATGCTTTAATTCAAAAAAGCTGCGCTTCAGGCAAAAAAGAGTCTTG 95
 107 TGCCCAAGATGCTTTAATTCAAAAAAGCTGCGCTTCAGGCAAAAAAGAGTCTTG 166
 96 GTAAATGCTGTCCTTAAATTCGCGCCACGCTTAATAGACACAAAGTAGAGGTGC 155
 167 GTAAATGCTGTCCTTAAATTCGCGCCACGCTTAATAGACACAAAGTAGAGGTGC 226
 156 TGAGTAGCTCTACAGAGTACAGAGGAGTACACCAAAACAAGAGGAGGAGAGA 215
 227 TGAGTAGCTCTACAGAGTACAGAGGAGTACACCAAAACAAGAGGAGGAGAGA 286
 216 TCATCAAGAACCTCATTAAGACAGTCTCAAGCTGCGCATTTTATAGAAATATCAGT 275
 287 TCATCAAGAACCTCATTAAGACAGTCTCAAGCTGCGCATTTTATAGAAATATCAGT 346
 276 TTATCAAGATGAGCTGAGATTGATGAGAAATTTAAGAGAAAGTTCATCAGTCTGA 335
 347 TTATCAAGATGAGCTGAGATTGATGAGAAATTTAAGAGAAAGTTCATCAGTCTGA 406
 336 TGAACGCTGCTCACTTCATCAAGTGTGATTAATCTTTGACCGAATGTGTTATCCAGC 395
 407 TGAACGCTGCTCACTTCATCAAGTGTGATTAATCTTTGACCGAATGTGTTATCCAGC 466
 396 TGTTAATGATGAGAGAGATGCTGACCAAAATCATTCAGCGCCACCTCACTGCCAAGT 455

Db 467 TGTTAATGATGAGAGAGATGCTGACCAAAATCATTCAGCGCCACCTACTGCCAAGT 526
 456 CACATGAGCGGTTAATTAATGCTTTCATCATTTTCAGATTGGAATTTTGGCTGCT 515
 527 CACATGAGCGGTTAATTAATGCTTTCATCATTTTCAGATTGGAATTTTGGCTGCT 566
 516 TGTTAATGCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 575
 587 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATGATCAAC 646
 576 AATGTTGATGAGAGACATAT 599
 647 AATGTTGATGAGAGACATAT 670

RESULT 29

LOCUS CD701806 622 bp mRNA linear EST 25-JUN-2003
 DEFINITION EST18330 human nasopharynx Homo sapiens cDNA, mRNA sequence.
 ACCESSION CD701806
 VERSION CD701806.1 GI:32232436
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 622)
 AUTHORS Liu X.-Q., Zhou Y., Zhang L.-J., Xu H., Chen H.-K., Pan Z.-G. and
 Transcriptional Gene Expression Profile of Human Nasopharynx
 Unpublished (2003)
 CONTACT Yixin Zeng
 Cancer Center
 Sun Yat-sen University
 651 Dongfeng Road East, Guangzhou 510060, China
 Tel: 86-1380-9770-743
 Fax: 86-20-8775-4506
 Email: yxzeng@zsums.edu.cn

FEATURES

source

1. 622
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="normal nasopharynx"
 /clone_id="human nasopharynx"
 /note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN

Query Match 89.4%; Score 537; DB 6; Length 622;
 Best Local Similarity 100.0%; Pred. No. 7e-138;
 Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

36 TGCCCAAGATGCTTTAATTCAAAAAAGCTGCGCTTCAGGCAAAAAAGAGTCTTG 95
 86 TGCCCAAGATGCTTTAATTCAAAAAAGCTGCGCTTCAGGCAAAAAAGAGTCTTG 145
 96 GTAAATGCTGTCCTTAAATTCGCGCCACGCTTAATAGACACAAAGTAGAGGTGC 155
 146 GTAAATGCTGTCCTTAAATTCGCGCCACGCTTAATAGACACAAAGTAGAGGTGC 205
 156 TGAGTAGCTCTACAGAGTACAGAGGAGTACACCAAAACAAGAGGAGGAGAGA 215
 206 TGAGTAGCTCTACAGAGTACAGAGGAGTACACCAAAACAAGAGGAGGAGAGA 265
 216 TCATCAAGAACCTCATTAAGACAGTCTCAAGCTGCGCATTTTATAGAAATATCAGT 275
 266 TCATCAAGAACCTCATTAAGACAGTCTCAAGCTGCGCATTTTATAGAAATATCAGT 325
 276 TTATCAAGATGAGCTGAGATTGATGAGAAATTTAAGAGAAAGTTCATCAGTCTGA 335
 326 TTATCAAGATGAGCTGAGATTGATGAGAAATTTAAGAGAAAGTTCATCAGTCTGA 385

QY	336	GGACCGNAGTCAGTTTCCATCAGAGTGATTTTAACTTTGGCCGGAAATGTATCCAGC	395
Db	386	TGACCGTGTCAgTTTCCATCAACGGGATATACCTTTGACCGGAATGTATTCACAGC	445
QY	396	TGTTAAATGAATGCAGAGAGATGTCTGCACCAAAATCATTACAGGCCACTCATCTGCCAGT	455
Db	446	TGTTAAATGAATGCAGAGAGATGTCTGCACCAAAATCATTACAGGCCACTCATCTGCCAGT	505
QY	456	CACATGACGCGTTATATATATGCTTTGATCATTTTCAGATTGTGAATTTTGGCTGCT	515
Db	506	CACATGACGCGTTATATATATGCTTTGATCATTTTCAGATTGTGAATTTTGGCTGCTGCT	565
QY	516	TGTATAATCCCTTTTGGGAATTTTAAACCCACTTACAAAATCTATGTATGGTATCA	572
Db	566	TGTATAATCCCTTTTGGGAATTTTAAACCCACTTACAAAATCTATGTATGGTATCA	622
RESULT 30			
BP228279			
LOCUS	BP228279	580 bp	mRNA linear EST 15-SEP-2004
DEFINITION	BP228279 Sugano cDNA library, dermoid cancer Homo sapiens cDNA		
ACCESSION	BP228279		
VERSION	BP228279.1	GI:52101189	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukacynska; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 580)		
JOURNAL	Suzuki,Y., Yamashita,R., Shiroya,M., Sakakibara,Y., Chiba,J.,		
COMMENT	Mizushima-Sugano,J., Nakai,K. and Sugano,S.		
	Sequence comparison of human and mouse genes reveals a homologous		
	block structure in the promoter regions		
	Genome Res. 14 (9), 1711-1718 (2004)		
	Contact: Yutaka Suzuki		
	Department of Virology		
	Institute of Medical Science, University of Tokyo		
	4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan		
	Email: yusuzuki@ims.u-tokyo.ac.jp.		
FEATURES	Location/Qualifiers		
source	1..580		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DMC08945"		
	/clone_lib="Sugano cDNA library, dermoid cancer"		
	/note="dermoid cancer"		
ORIGIN			
Query Match	89.0%;	Score 535;	DB 5;
Beet Local Similarity	99.6%;	Pred.No.2.5e-137;	Length 580;
Matches	546;	Conservative	0;
		Mismatches	1;
		Indels	1;
		Gaps	14;
QY	54	ATTCCAAAAACCTGCGCCGTTCAAGCACAAAAGAAATCTGGGTAAATGATGTCCTCAAT	113
Db	1	ATTCCAAAAACCTGCGCCGTTCAAGCACAAAAGAAATCTGGGT- AATGATGTCCTCAAT	59
QY	114	CCATGCGCACCACTTTAATAGACACACAGTATGATGAGTGTGGATGAGCTTACAGAG	173
Db	60	CCATGCGCACCACTTTAATAGACACACAGTATGATGAGTGTGGATGAGCTTACAGAG	119
QY	174	TGACCAAGGAGATGACCCAAAACAAAGAGAGAGAGAGATGATCAAGAAACCTCATCA	233
Db	120	TGACCAAGGAGATGACCCAAAACAAAGAGAGAGAGAGATGATCAAGAAACCTCATCA	179
QY	234	AGACAGTCATCAAGCTGGCCATCTTTTATAGGAATATCAGTTTATCAAGATGAGCTAG	293
Db	180	AGACAGTCATCAAGCTGGCCATCTTTTATAGGAATATCAGTTTATCAAGATGAGCTAG	239
QY	294	CATTGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCTATGACCGTGTCACTTTC	353

Df	240	CATTGATGGAGAAATTTAAGAAGAAAGTTTATCATCAGCTTGTAATGACCGGTGCAGTTTC	299
Oy	354	ATCAGGTGGAATTAATTAACCTTTGACCCGGAATGTTTATTCACAGGCTGTTAATGATGCAGAG	413
Df	300	ATCAGGTGGAATTAATTAACCTTTGACCCGGAATGTTTATTCACAGGCTGTTAATGATGCAGAG	359
Oy	414	AGATGCTGCAACCAAATTAATTCAGGCGCACCTCACTGCCAAGTACAATGACAGGAGTTATA	473
Df	360	AGATGCTGCAACCAAATTAATTCAGGCGCACCTCACTGCCAAGTACAATGACAGGAGTTATA	419
Oy	474	ATGCTTTGATCATTTTTTCAGATTTGATGAATTTTTTGCGCTTGTATAATCCTTTTGGGA	533
Df	420	ATGCTTTGATCATTTTTTCAGATTTGATGAATTTTTTGCGCTTGTATAATCCTTTTGGGA	479
Oy	534	ATTTTAAACCCCACTTAACAAAAACTATGTGATGTATCAACAAAATGTTGGATGACAGA	593
Df	480	ATTTTAAACCCCACTTAACAAAAACTATGTGATGTATCAACAAAATGTTGGATGACAGA	539
Oy	594	ACATATGA 601	
Df	540	ACATATGA 547	
<hr/>			
RESULT 31			
LOCUS	CN305328		
DEFINITION	17000599947340 GRN PRENEU Homo sapiens cDNA 5', mRNA sequence.	672 bp	mRNA linear EST 16-MAY-2004
ACCESSION	CN305328		
VERSION	CN305328.1	GI:47321742	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 672) Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J and Stanton, L.W. Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation Nat. Biotechnol. 22 (6), 707-716 (2004) Contact: Brandenberger R Regenerative Medicine Genon Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@genon.com Insert Length: 672 Std Error: 0.00.		
JOURNAL COMMENT	Location/Qualifiers		
TITLE	1..672		
FEATURES	/organism="Homo sapiens"		
source	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"		
	/clone_lib="GRN PRENEU"		
	/note="oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic acid and mitogens."		
ORIGIN			
Query Match	88.8%; Score 533.6; DB 7; Length 672;		
Bases Similarity	98.8%; Pred. No. 6.3e-117;		
Matches 569; Conservative	0; Mismatches 4; Indels 3; Gaps 3;		
Oy	1	CTTCAGGCTCCCGAGCGCGTGC-CGACTCTTCGATGAGCACAGATGCTTTAATCCA	59
Df	97	CTTAGAGTCCCGGCGCGTGCCTCTCCGATGAGCACAGATGCTTTAATCCA	156
Oy	60	AAAACCTGGCGTTCAAGCAAAAAGATCTTGGTAAATGTTGTCCAAATCATCG	119

Db 157 AAAAAGCTGGCGTTGAGGACAAAGAAAGATCTTGGTAAATGTGTCCAAATCCATCG 216

Qy 120 CCACCACTTAATATGACAGACAAAGTATGAGGTGCTGATGAGCTCTTACAGAGTGACCA 179

Db 217 CCACCACTTAATATGACAGACAAAGTATGAGGTGCTGATGAGCTCTTACAGAGTGACCA 276

Qy 180 GGGAGTACACCCCAAAACAAAGAGGAGGAGAGAGATCATCAGAACCTTCAAGACAG 239

Db 277 GGGAGTACACCCCAAAACAAAGAGGAGGAGAGAGATCATCAGAACCTTCAAGACAG 336

Qy 240 TCATCAAGCTGGCCATCTTATAGAAATATCATGTTAATCAAGATGAGTACATTTGA 299

Db 337 TCATCAAGCTGGCCATCTTATAGAAATATCATGTTAATCAAGATGAGTACATTTGA 396

Qy 300 TGAAGAAATTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGATGATTCATGAG 359

Db 397 TGAAGAAATTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGATGATTCATGAG 456

Qy 360 TGAATTAATACCTTTGACCGGAATGTCTTATCCAGGCTTTAAATGAATGACAGAGATGC 419

Db 457 TGAATTAATACCTTTGACCGGAATGTCTTATCCAGGCTTTAAATGAATGACAGAGATGC 516

Qy 420 TGCACCAATCATTTACAGCGCCACCTCACTGCAAGTACAT- GAGCGGTTAATTAATGTC 478

Db 517 TGCACCAATCATTTACAGCGCCACCTCACTGCAAGTACATGAGAGCGGTTAATTAATGTC 576

Qy 479 TTTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAAT- CTTTGGGAATTT 537

Db 577 GTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTT 636

Qy 538 TAAACCCCACTTACAAAACATATGTATGTATCA 573

Db 637 TAAACCCCACTTACAAAACATATGTATGTATCA 672

RESULT 32
LOCUS CV023086 566 bp mRNA linear EST 20-AUG-2004

DEFINITION 5463 Full length cDNA from the Mammalian Gene Collection Homo sapiens cDNA 5' similar to BC007014, mRNA sequence.

ACCESSION CV023086

VERSION CV023086.1 GI:51480836

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE AUTHORS Mammalia: Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 566)

TITLE Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Driscoll, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O., Clindingsmith, T.R., Hartley, J.L., Espósito, D., Cheo, D., Moore, T., Simons, B., Segueria, R., Bosak, S., Doucette-Stamm, L., Le Pench, C., Vandenhaute, J., Quisic, M.E., Alcala, J.S., Hill, D.E. and Vidal, M. Human ORFome Version 1.1: a Platform for Reverse Proteomics Genome Res. (2004) In press

JOURNAL COMMENT

CONTACT: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF results from a PCR reaction using an MGC full-length cDNA as template DNA and ORF specific primers
PCR Primers
FORWARD: ATGCACTCCGAGCAGACAGAAATCCAA
BACKWARD: CATATGTTCTCTTCAATCCAAATTT
Insert Length: 566 Std Error: 21.00
Plate: 11075 row: 05 column: F
Seq primer: ACTGGCGTGTCTTACAAACGCTGTGACGTGGAAGAAC
High quality sequence start: 104
High quality sequence stop: 565

FEATURES POLYA=No. Location/Qualifiers

source 1..566

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="mixed"

/clone_lib="Full Length cDNA from the Mammalian Gene Collection"

/note="Vector: mixed; The ORFs were PCR amplified from the MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational Gateway cloning into pDONR223 Donor Vector. Reference: MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-length Human and Mouse cDNA Sequences. PNAS, 2002, 99(126), 16899-16903"

Query Match 88.8%; Score 533.4; DB 7; Length 566;
Best Local Similarity 99.8%; Pred. No. 6.8e-137;
Matches 534; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 36 TGGCCACAGATGCTTAAATTCAAAACCTGGCCGTCGAGCAGAAAAGATCTTG 95

Db 32 TGGCCACAGATGCTTAAATTCAAAACCTGGCCGTCGAGCAGAAAAGATCTTG 91

Qy 96 GTAAATGTGTCCAAATCCATGCGCACCACTTAATAGACACAAAGTATGAGGTGC 155

Db 92 GTAAATGTGTCCAAATCCATGCGCACCACTTAATAGACACAAAGTATGAGGTGC 151

Qy 156 TGGATGAGCTTACAGAGTACAGAGAGTACACCCAAAACAAAGAGAGAGAGAGA 215

Db 152 TGGATGAGCTTACAGAGTACAGAGAGTACACCCAAAACAAAGAGAGAGAGA 211

Qy 216 TCATCAAGAACCTTCAAGACAGATGATGAGTGCATCTTTATAGAAATATCACT 275

Db 212 TCATCAAGAACCTTCAAGACAGATGATGAGTGCATCTTTATAGAAATATCACT 271

Qy 276 TTAATCAAGTATGAGTATGATGATGAGAGAAATTTAAGAAAGTATCATCTTGCTA 335

Db 272 TTAATCAAGTATGAGTATGATGATGAGAGAAATTTAAGAAAGTATCATCTTGCTA 331

Qy 336 TGAACGCTGCTCACTTCCATCAGGTGATTAATCTTTGACCGGAATGTATTCAGGC 395

Db 332 TGAACGCTGCTCACTTCCATCAGGTGATTAATCTTTGACCGGAATGTATTCAGGC 391

Qy 396 TGTTAATGAATGACAGAGATGCTGACCAATCATTTACAGCGCCACCTTACCTCAAGT 455

Db 392 TGTTAATGAATGACAGAGATGCTGACCAATCATTTACAGCGCCACCTTACCTCAAGT 451

Qy 456 CACATGAGAGGGTTAATATGCTTTGATCATTTTTCAGATTTGGAATTTTGGCTGCT 515

Db 452 CACATGAGAGGGTTAATATGCTTTGATCATTTTTCAGATTTGGAATTTTGGCTGCT 511

Qy 516 TGTATATCTTTTGGAAATTTTAAACCCCACTTACAAAACATATGATGATGAT 570

Db 512 TGTATATCTTTTGGAAATTTTAAACCCCACTTACAAAACATATGATGATGAT 566

RESULT 33

LOCUS CR557522

DEFINITION DKFZp469E2119.1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone

ACCESSION CR557522

VERSION CR557522.1 GI:50278801

KEYWORDS EST.

SOURCE Pongo pygmaeus (orangutan)

ORGANISM Pongo pygmaeus

REFERENCE AUTHORS Mammalia: Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Pongo. 1 (bases 1 to 707)
Pouetka, A., Albert, R., Moosmayer, P., Schnupp, I., Wellenreuther, R.,

TITLE Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobbo,G., Han,M. and Wiemann,S.
 JOURNAL Pongo pygmaeus mRNA (Poustka,A., Albert,R., Moosmayer,P., et al.)
 COMMENT Unpublished (2004)
 CONTACT: MIPS

MIPS Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
 This is the 5' sequence of the clone insert from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp4982119
 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
 source location/Qualifiers

1..707
 /organism="Pongo pygmaeus"
 /mol_type="mRNA"
 /db_xref="taxon:9600"
 /clone="DKFZp4982119"
 /tissue_type="kidney"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1ib="469 (synonym: pkid1)"
 /note="Vector: pSPORT1_sfi; site_1: sfi1A; site_2: sfi1B"

ORIGIN

Query Match 88.3%; Score 530.6; DB 7; Length 707;
 Best Local Similarity 99.3%; Pred. No. 4,3e-136;
 Matches 533; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTAATTCGAAACCTGGCGGTGACGACAAAAGATCTTG 95

DB 171 TGGCCACAGATGCTTAATTCGAAACCTGGCGGTGACGACAAAAGATCTTG 230

QY 96 GTAAATGCTGTCGAATTCATGCCACACCTTAATAGACACAGATGAGTGC 155

DB 231 GTAAATGCTGTCGAATTCATGCCACACCTTAATAGACACAGATGAGTGC 290

QY 156 TGGAGAGCTGTCACAGAGTACGACGAGATGACCCAAACAGAGAGGAGAGAA 215

DB 291 TGGAGAGCTGTCACAGAGTACGACGAGATGACCCAAACAGAGAGGAGAGAA 350

QY 216 TCATCAGAACCTCATCAGACAGTCAAGTGCCTTTTATAGGAATATCACT 275

DB 351 TCATCAGAACCTCATCAGACAGTCAAGTGCCTTTTATAGGAATATCACT 410

QY 276 TTAATCAGATGAGTCAATGATGAGAGAAATTTAAGAGAAAGTTCACTGCT 335

DB 411 TTAATCAGATGAGTCAATGATGAGAGAAATTTAAGAGAAAGTTCACTGCT 470

QY 336 TGAACCTGCTCACTTCCATCAGTGTGATTTATCCTTGAACCGGATGTTTCAAGC 395

DB 471 TGAACCTGCTCACTTCCATCAGTGTGATTTATCCTTGAACCGGATGTTTCAAGC 530

QY 396 TGTAAATGAATGACAGAGATGCTGACCAATCATTCAGAGCGACCTCACTGCCAAGT 455

DB 531 TGTAAATGAATGACAGAGATGCTGACCAATCATTCAGAGCGACCTCACTGCCAAGT 590

QY 456 CACATGACGGGTAAATTAATGCTTGTATCAATTTTCAATGTTGAATTTTGGTGCCT 515

DB 591 CACATGACGGGTAAATTAATGCTTGTATCAATTTTCAATGTTGAATTTTGGTGCCT 650

QY 516 TGTATTAATCTTTTGGGAATTTAAACCCCATTAACAAAACCTATGTATGATCA 572

DB 651 TGTATTAATCTTTTGGGAATTTAAACCCCATTAACAAAACCTATGTATGATCA 707

RESULT 34

CK903031 572 bp mRNA linear EST 11-MAR-2004
 LOCUS CK903031
 DEFINITION 1155c03.x5 HR85 islet Homo sapiens cDNA clone IMAGE:6548934 3'
 similar to TR:095379 095379 MDC-3.13 ISOFORM 2. [2] TR:Q9UP47 ;

ACCESSION CK903031
 VERSION CK903031.1 GI:45364562
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 572)

Wyllie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,D., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarelisvili,R., Williams,T., Jackson,Y. and Bowers,Y.
 WashU-Harvard Pancreas EST Project

Unpublished (2000)
 Other ESTs: 1155c03.y1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

This read is a 3' RESEQUENCE of a previously sequenced pancreas clone

FEATURES

source

1..572
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6548934"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_1ib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1: NotI; Site 2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size -1kb. 5' xhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Reimut Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 88.2%; Score 529.8; DB 7; Length 572;
 Best Local Similarity 99.3%; Pred. No. 6,8e-136;
 Matches 553; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 CTTCAGCGTCCGGGCGCGTGC-CGACTCCCTCCGATGGCCACAGATGCTTTAATTC 59

DB 17 CTTCAGCGTCCGGGCGCGTGC-CGACTCCCTCCGATGGCCACAGATGCTTTAATTC 76

QY 60 AAAACCTGGCGCTTCAAGGACAAAAGATCTTGGTAAATGTTGTTCAATCCATCG 119

DB 77 AAAACCTGGCGCTTCAAGGACAAAAGATCTTGGTAAATGTTGTTCAATCCATCG 136

QY 120 CCACCACTTAATAGACGACAAAGTGTGAGGTGCTGATAGAGCTCTACAGAGTACCA 179

DB 137 CCACCACTTAATAGACGACAAAGTGTGAGGTGCTGATAGAGCTCTACAGAGTACCA 196

QY 180 GGGAGTACCCCAAAACAAAGAGAGGCGAGAGATCATCAAGAACCTCATCAAGACAG 239

DB 197 GGGAGTACCCCAAAACAAAGAGAGGCGAGAGATCATCAAGAACCTCATCAAGACAG 256

QY 240 TCATCAGCTGCGCATCTTTTATAGAAATATCAATTAATCAAGATGAGCTAGCATTTGA 299

Db 257 TCATCAAGCTGGGCATTCTTATAGAAATATCATGTTTATCAAGATGACGTAGCAATGA 316
QY 300 TGGAGAAATTTAAGAAAGAGTTTCATCAGCTTGCTATGACCGGTGTCAGTTTCATCAGG 359
Db 317 TGGAGAAATTTAAGAAAGAGTTTCATCAGCTTGCTATGACCGGTGTCAGTTTCATCAGG 376
QY 360 TGGATTATACCTTTGACCGGAAATGTGTATTCAGAGCGGCTTTAAATGAATGAGAGACATGC 419
Db 377 TGGATTATACCTTTGACCGGAAATGTGTATTCAGAGCGGCTTTAAATGAATGAGAGATGC 436
QY 420 TGCACCAATCATTCAGCGCCACCTCATCTGCAAGTCACATGAGCGGTTAAATATGCT 479
Db 437 TGCACCAATCATTCAGCGCCACCTCATCTGCAAGTCACATGAGCGGTTAAATATGCT 496
QY 480 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTGTAATTCCTTTTGGGAATTTTA 539
Db 497 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTGTAATTCCTTTTGGGAATTTTA 555
QY 540 AACCCCACTTACAAAA 556
Db 556 AACCCCACTTACAAAA 572

RESULT 35
AV713586 697 bp mRNA linear EST 11-OCT-2000
LOCUS AV713586 DCB Homo sapiens cDNA clone DCBBB11 5', mRNA sequence.
DEFINITION AV713586
ACCESSION AV713586
VERSION AV713586.1 GI:10795103
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Xu, X., Gu, J., Liu, F., Qu, J., Zhao, M., Li, Y., Huang, Q., Zhou, J.,
Song, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X.,
Cheng, Z., Xu, S., Yu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,
Lu, G., Cheng, Z., and Han, Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

TITLE JOURNAL
COMMENT

FEATURES
source
1. 697
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBBB11"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_1ib="DCB"
/note="Vector: pTriplex2; Site_1: sf11a; Site_2: sf11b"

ORIGIN
Query Match 88.1%; Score 529.4; DB 1; Length 697;
Best Local Similarity 98.9%; Pred. No. 9.3e-136;
Matches 533; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTGAGGACACAAAGAAATCTTGG 95
Db 159 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTGAGGACACAAAGAAATCTTGG 218
QY 96 GTAAATGCTGTCAGATCGCACCACTTAAATGAGACACAGAAGTAGAGGTGC 155
Db 219 GTAAATGCTGTCAGATCGCACCACTTAAATGAGACACAGAAGTAGAGGTGC 278

QY 156 TGGATAGCTTACAGATGACGAGGAGTACACCCAAAAAAGAGAGGCGAGAGAGA 215
Db 279 TGGATAGCTTACAGATGACGAGGAGTACACCCAAAAAAGAGAGGCGAGAGAGA 338
QY 216 TCATCAAGAACCTCATCAAGACAGCATCAAGCTGGCCATTCTTTATAGGAATTAACGT 275
Db 339 TCATCAAGAACCTCATCAAGACAGCATCAAGCTGGCCATTCTTTATAGGAATTAACGT 398
QY 276 TTAATCAAGATGAGTACGATGATGAGAGAAATTTAAGAAAGAGTTCACTGCTGA 335
Db 399 TTAATCAAGATGAGTACGATGATGAGAGAAATTTAAGAAAGAGTTCACTGCTGA 458
QY 336 TGACCGTGTGCTGCTTCATCAGTGCGATTATACCTTTGACCGGAATGTTATCAAGC 395
Db 459 TGACCGTGTGCTGCTTCATCAGTGCGATTATACCTTTGACCGGAATGTTATCAAGC 518
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATGAGCGCCACTCAGCCCAAGT 455
Db 519 TGTAAATGAATGACAGAGATGCTGCACCAATTCATGAGCGCCACTCAGCCCAAGT 578
QY 456 CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
Db 579 CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 638
QY 516 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAAAATGATGTATCAAC 574
Db 639 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAAAATGATGTATCAAC 697

RESULT 36
BG434534 739 bp mRNA linear EST 14-MAR-2001
LOCUS BG434534
DEFINITION 602506586F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:460317 5',
mRNA sequence.
ACCESSION BG434534
KEYWORDS BG434534.1 GI:13341040
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CLOUTech Laboratories, Inc.
cDNA Library Preparation: CLOUTech Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1338 row: k column: 22
High quality sequence stop: 717.
Location/Qualifiers
1. 739
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:460317"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NIH_MGC_79"
/note="Torgan: Placenta; Vector: pDMR-LIB (Clontech);
Site_1: Sf11 (ggccggccggcc); Site_2: Sf11
(ggccatcatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCCATTTATGCCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

ORIGIN

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

Query Match

87.8%; Score 527.6; DB 4; Length 739;
Best Local Similarity 97.7%; Pred. No. 3e-135;

Matches 588; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

```

QY 1 CTTGACGCTCCCGGCGCGCTGCG-CGACTCTCTCCGATGCGCCACAGATGCTTTAATCCA 59
    |||
DB 76 CTTGACGCTCCCGGCGCGCTGCGCGCCTCTCCGATGCGCCACAGATGCTTTAATCCA 135
    |||

QY 60 AAAACCTGCGCGCTTTCAGGACCAAAAAGATCTTGGGTAAATGGGTCCAAATTCATCG 119
    |||
DB 136 AAAACCTGCG-CGTTTCAGGACCAAAAAGATCTTGGGTAAATGGGTCCAAATTCATCG 194
    |||

QY 120 CCACCACTTAATAGACGACACAGATGAGGTGCTGAGTGAAGCTCTACAGAGTACCA 179
    |||
DB 195 CCACCACTTAATAGACGACACAGATGAGGTGCTGAGTGAAGCTCTACAGAGTACCA 254
    |||

QY 180 GGGAGTACACCCAAAACAAGAGAGGACAGAGATTCATCAAGAACCTCATCAAGACAG 239
    |||
DB 255 -GGAGTACACCCAAAACAAGAGAGGACAGAGATTCATCAAGAACCTCATCAAGACAG 313
    |||

QY 240 TCATCAAGCTGGCCATCTTTAATAGGAATATCATGTTAATCAAGATGAGCTAGCATTTGA 299
    |||
DB 314 TCATCAAGCTGGCCATCTTTAATAGGAATATCATGTTAATCAAGATGAGCTAGCATTTGA 373
    |||

QY 300 TGGAGAAATTTAAGAAAGAGTTGATCAGCTGCTATGACCGTGTACGTTTCATCAG 359
    |||
DB 374 TGGAGAAATTTAAGAAAGAGTTGATCAGCTGCTATGACCGTGTACGTTTCATCAG 433
    |||

QY 360 TGGATATACCTTTGACCGGAATGTTTATCCAGGCTGTTAATGTAATGACAGAGATGC 419
    |||
DB 434 TGGATATACCTTTGACCGGAATGTTTATCCAGGCTGTTAATGTAATGACAGAGATGC 493
    |||

QY 420 TGCACCAATCATCTTTCAGCGCCACCTCATGCGCAAGTCACATGAGCGGTTAATATGCT 479
    |||
DB 494 TGCACCAATCATCTTTCAGCGCCACCTCATGCGCAAGTCACATGAGCGGTTAATATGCT 553
    |||

QY 480 TTGATCATTTTTCAAGTGTGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTA 539
    |||
DB 554 TTGATCATTTTTCAAGTGTGAATTTTGGCTGCTG-TGTATTAATCCTTTGGGAATTTT 612
    |||

QY 540 AACCCCACTTACAAAACATATGATGATGATCACAAAATGTTGATGAAGAGAACTAT 599
    |||
DB 613 AACCCCACTTACAAAACATATGATGATGATCACAAAATGTTGATGAAGAGAACTAT 671
    |||

QY 600 GA 601
    ||
DB 672 GA 673
    ||

```

RESULT 37
CR71358 692 bp mRNA linear EST 23-SEP-2004

LOCUS DKEP46910535_r1_469 (synonym: pkid1) Pongo pygmaeus cDNA clone

DEFINITION DKEP46910535_5', mRNA sequence.

ACCESSION CR71358.1 GI:52614631

VERSION CR71358.1

KEYWORDS EST.

SOURCE Pongo pygmaeus (orangutan)

ORGANISM Pongo pygmaeus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.

1 (bases 1 to 692)

Pousetka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
Pongo pygmaeus mRNA (Pousetka, A., Albert, R., Moosmayer, P., et al.)
Unpublished (2004)
COMMENT Contact: MIPS

FEATURES

source

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for
ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKEP46910535
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

Location/Qualifiers

```

1..692
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKEP46910535"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="469 (synonym: pkid1)"
/note="Vector: pSPORT1_sfi; Site_1: sfi1a; Site_2: sfi1b"

```

ORIGIN

Query Match

87.6%; Score 526.6; DB 7; Length 692;
Best Local Similarity 99.1%; Pred. No. 5.5e-135;

Matches 540; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```

QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCGGTCAGGACACAAAAGAGATCTTGG 95
    |||
DB 149 TGGCCACAGATGCTTTAATTCAAAACCTGGCGGTCAGGACACAAAAGAGATCTTGG 208
    |||

QY 96 GTAAATGGTGTCCAATTCATGCGCACCTTAATAGACACACAAGATGAGGTGC 155
    |||
DB 209 GTAAATGGTGTCCAATTCATGCGCACCTTAATAGACACACAAGATGAGGTGC 268
    |||

QY 156 TGGATGAGCTTACAGAGTACAGAGTACACCCAAAACAAGAGAGGACAGAGA 215
    |||
DB 269 TGGATGAGCTTACAGAGTACAGAGTACACCCAAAACAAGAGAGGACAGAGA 328
    |||

QY 216 TCATCAAGAACCTCATCAAGACATGATCAGCTGGCCATCTTTAATAGGAATATCAGT 275
    |||
DB 329 TCATCAAGAACCTCATCAAGACATGATCAGCTGGCCATCTTTAATAGGAATATCAGT 388
    |||

QY 276 TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 335
    |||
DB 389 TTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 448
    |||

QY 336 TGAACGTGTCACTTTCATCAGTGTGATTAATCCTTGAACCGGAATGTTATCCAGGC 395
    |||
DB 449 TGAACGTGTCACTTTCATCAGTGTGATTAATCCTTGAACCGGAATGTTATCCAGGC 508
    |||

QY 396 TGTAAATGAATGACAGAGATGCTGACCAATCATTCAGGCGCACCTCATGCGCAAGT 455
    |||
DB 509 TGTAAATGAATGACAGAGATGCTGACCAATCATTCAGGCGCACCTCATGCGCAAGT 568
    |||

QY 456 CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATTTGATGAAATTTTGGCTGCT 515
    |||
DB 569 CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATTTTGGCTGCTGCT 628
    |||

QY 516 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATTAATGATGATCAACA 575
    |||
DB 629 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATTAATGATGATCAACA 687
    |||

QY 576 AATG 580
    ||||
DB 688 AATG 692
    ||||

```

RESULT 38

LOCUS BG534704

DEFINITION

ACCESSION

BG534704 685 bp mRNA linear EST 03-APR-2001
602553867F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:463672 5',
mRNA sequence.
BG534704


```

VERSION      BG534704.1  GI:13526246
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS      1 (bases 1 to 685)
TITLE        NIH-MGC http://mgc.nci.nih.gov/
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT       Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgaabs-remail.nih.gov
              Tissue Procurement: CLONTECH Laboratories, Inc.
              CDNA Library Preparation: CLONTECH Laboratories, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.jnl.gov
              Plate: LICM1466 row: e column: 17
              High quality sequence stop: 683.
              Location/Qualifiers
                1..685
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:463672"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site 1:
                5'fl (ggcgccctggcgc); Site 2: 5'fl (ggccatcatggc); 5' and
                3' adaptors were used in cloning as follows: 5' adaptor
                sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                5'-ATTCTAGAGCCGACGGCCGACATG-dt(30)BN-3' (where B = A,
                C, or G and N = A, C, G, or T). Average insert size 1.9
                kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                by PCR. This library was enriched for full-length clones
                and was constructed by Clontech Laboratories (Palo Alto,
                CA). Note: this is a NIH_MGC Library."
ORIGIN
Query Match      86.5%; Score 519.8; DB 4; Length 685;
Best Local Similarity 99.3%; Pred. No. 4,3e-133;
Matches 543; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 55 TTCCAAAAAAGCTGGCCCTTTCAGGACCAAAAGAGATCTTGGTAAATGGTGTCCAATC 114
Db 1 TTCCAAAAAAGCTGGCCCTTTCAGGACCAAAAGAGATCTTGGTAAATGGTGTCCAATC 60
QY 115 CATGCCACCACTTAAATAGACGACACAGTAGTAGAGTCTGGATGAGCTTACAGAGT 174
Db 61 CATGCCACCACTTAAATAGACGACACAGTAGTAGAGTCTGGATGAGCTTACAGAGT 120
QY 175 GACCAAGGAGTACACCCAAAGAGAGGAGGAGAGAAATCATCAAGAACTCATCA 224
Db 121 GACCA-AGAGTAAACCCAAAGAGAGGAGGAGAGAAATCATCAAGAACTCATCA 179
QY 235 GACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAGTTTAAATCAAGAGCTGAC 294
Db 180 GACAGTCATCAAGCTGGCCATCTTTATAGAAATATCAGTTTAAATCAAGAGCTGAC 239
QY 295 ATTGATGAGAAATTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGCTAGTTTCA 354
Db 240 ATTGATGAGAAATTTAAGAGAAAGTTTCATCAGCTTGTATGACCGTGTGCTAGTTTCA 299
QY 355 TCAGGTGAGATTATACCTTTTGAACGGGAATGTGTATTCAGGCTGTAAATGAATGCAGAGA 414
Db 300 TCAGGTGAGATTATACCTTTTGAACGGGAATGTGTATTCAGGCTGTAAATGAATGCAGAGA 359
QY 415 GATGCTGACCAATATCATTCAGGCGCCACCTCATGCAATCATCATGAGAGGGTTAATA 474
Db 360 GATGCTGACCAATATCATTCAGGCGCCACCTCATGCAATCATCATGAGAGGGTTAATA 419

```

```

QY 475 TGCTTTGATCATTTTTCAGATTGTGAATTTTGGCGCTTGTATATCCTTTGGGAA 534
Db 420 TGCTTTGATCATTTTTCAGATTGTGA-TTTTGGCGCTTGTATATCCTTTGGGAA 478
QY 535 TTTTAAACCCCACTTCAAAAACCTATGATGTGTATCAACAAATGTTGATGAGAGAA 594
Db 479 TTTTAAACCCCACTTCAAAAACCTATGATGTGTATCAACAAATGTTGATGAGAGAA 538
QY 595 CATATGA 601
Db 539 CATATGA 545
RESULT 39
LOCUS      CV027255
DEFINITION 5464 Full Length cDNA from the Mammalian Gene Collection Homo
              sapiens CDNA 5' similar to BC005352, mRNA sequence.
ACCESSION  CV027255
VERSION     CV027255.1  GI:51485291
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
              1 (bases 1 to 593)
              Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
              Driscoll,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
              Clingingsmith,T.R., Hartley,J.L., Espósito,D., Cheo,D., Moore,T.,
              Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
              Vandenhoute,J., Chisick,M.E., Albalade,J.S., Hill,D.E. and Vidal,M.
              Human ORFome Version 1.1: a Platform for Reverse Proteomics
              Genome Res. (2004) In press
              Contact: Vidal M
              Marc Vidal Laboratory
              Dana Farber Cancer Institute
              1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
              Tel: 617 632 5180
              Fax: 617 632 5739
              Email: Marc.Vidal@dfci.harvard.edu
              ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
              results from a PCR reaction using an MGC full-length cDNA as
              template DNA and ORF specific primers
              PCR Primers
              FORWARD: ATGCACCTCCAGACAGAAAGATCCAA
              BACKWARD: CATATGCTCTTCATCCAACTTT
              Insert Length: 593 Std Error: 78.00
              Plate: 11028 row: 11 column: H
              Seq primer: ACTGGCCGTGTTTACAAAGTCGTGACTGGGAAAC
              High quality sequence start: 101
              High quality sequence stop: 592
              POLVA=No.
              Location/Qualifiers
                1..593
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /tissue_type="mixed"
                /clone_lib="Full length cDNA from the Mammalian Gene
                Collection"
                /note="Vector: mixed; The ORFs were PCR amplified from the
                MGC (Mammalian Gene Collection) as of April 2004 and
                cloned by recombinational Gateway cloning into pDONR23
                donor vector. Reference: MGC (Mammalian Gene Collection)
                Program Team, Generation and Initial Analysis of more than
                15,000 full-length Human and Mouse cDNA Sequences. PNAS,
                2002, 99(126), 16899-16903"
ORIGIN
Query Match      86.4%; Score 519.2; DB 7; Length 593;
Best Local Similarity 98.0%; Pred. No. 6,1e-133;
Matches 547; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

```

```
QY 36 TGGCCACAGATGCTTTTAATTCACAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTGG 95
DB 32 TGGCCACAGATGCTTTTAATTCACAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTGG 91
QY 96 GTTAAATGCTGTCCTTAATTCATCGCCACCACTTAATGACGACACAGTAGTAGTGC 155
DB 92 GTTAAATGCTGTCCTTAATTCATCGCCACCACTTAATGACGACACAGTAGTAGTGC 151
QY 156 TGGATGAGCTTACAGAGTACAGGAGGATGACACCCAAAACAAGAGAGGACAGAGA 215
DB 152 TGGACAGCTTACAGAGTACAGGAGGATGACACCCAAAACAAGAGAGGACAGAGA 211
QY 216 TCATCAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGAAATATCACT 275
DB 212 TCATCAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGAAATATCACT 271
QY 276 TTAATCAGATGAGTCAAGTATGATGAGAGAAATTTAAGAAAGAGTTCACTGCTTCTA 335
DB 272 TTAATCAGATGAGTCAAGTATGATGAGAGAAATTTAAGAAAGAGTTCACTGCTTCTA 331
QY 336 TGACCGTGTGCTGATTCATCAGTGGATTAATACCTTTGACCGGAATGTTATCCAGGC 395
DB 332 TGACCGTGTGCTGATTCATCAGTGGATTAATACCTTTGACCGGAATGTTATCCAGGC 391
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCACTGCCAAT 455
DB 392 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCACTGCCAAT 451
QY 456 CACATGACGCGTTAATTAATGCTTTGATCATTTTCAATGTTGAAATTTTGGCTGCT 515
DB 452 CACATGACGCGTTAATTAATGCTTTGATCATTTTCAATGTTGAAATTTTGGCTGCT 511
QY 516 TGTAAATGCTTTGGGAAATTTTAAACCCACTTCAAAAACATGATGATGATCAAA 575
DB 512 TGTAAATGCTTTGGGAAATTTTAAACCCACTTCAAAAACATGATGATGATCAAA 568
QY 576 AAATGTTGATGAGAGA 593
DB 569 AAATGTTGATGAGAGA 586

RESULT 40
LOCUS CD520319 757 bp mRNA linear EST 06-JUN-2003
DEFINITION AGSCOURT 14360077 NIH MGC 191 Homo sapiens cDNA clone
IMAGE:30410454 5', mRNA sequence.
ACCESSION CD520319
VERSION CD520319.1 GI:31452037
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 757)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: dsgerha@imail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: NDCM200 row: k column: 07
High quality sequence stop: 563.
Location/Qualifiers
1..757
FEATURES
source
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30410454"
/issue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NIH MGC 191"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggccgctggcc); Library is oligo-dT primed
and directionally cloned. PMMC - peripheral blood
mononuclear cells. RNA was pooled from 3/6hour stimulation
with PMA adn ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-ATTCTAAGAGCGGCGGCGGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC library."
```

ORIGIN

```
Query Match 85.5%; Score 513.6; DB 6; Length 757;
Best Local Similarity 98.7%; Pred. No. 2.3e-131;
Matches 539; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 36 TGGCCACAGATGCTTTTAATTCACAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTGG 95
DB 107 TGGCCACAGATGCTTTTAATTCACAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTGG 166
QY 96 GTTAAATGCTGTCCTTAATTCATCGCCACCACTTAATGACGACACAGTAGTAGTGC 155
DB 167 GTTAAATGCTGTCCTTAATTCATCGCCACCACTTAATGACGACACAGTAGTAGTGC 226
QY 156 TGGATGAGCTTACAGAGTACAGGAGTACCCAAAACAAGAGAGGACAGAGA 215
DB 227 TGGATGAGCTTACAGAGTACAGGAGTACCCAAAACAAGAGAGGACAGAGA 286
QY 216 TCATCAGAACCTCATCAAGACAGTCAAGCTGGCACTTTTATAGAAATATCACT 275
DB 287 TCATCAGAACCTCATCAAGACAGTCAAGCTGGCACTTTTATAGAAATATCACT 346
QY 276 TTAATCAGATGAGTCAAGTATGATGAGAAATTTAAGAAAGAGTTCACTGCTTCTA 335
DB 347 TTAATCAGATGAGTCAAGTATGATGAGAAATTTAAGAAAGAGTTCACTGCTTCTA 406
QY 336 TGACCGTGTGCTGATTCATCAGTGGATTAATACCTTTGACCGGAATGTTATCCAGGC 395
DB 407 TGACCGTGTGCTGATTCATCAGTGGATTAATACCTTTGACCGGAATGTTATCCAGGC 466
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCACTGCCAAT 455
DB 467 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCACTGCCAAT 526
QY 456 CACATGACGCGTTAATTAATGCTTTGATCATTTTTCAGATTTGAAATTTTGGCTGCT 515
DB 527 CACATGACGCGTTAATTAATGCTTTGATCATTTTTCAGATTTGAAATTTTGGCTGCT 586
QY 516 TGTAAATGCTTTGGG--AATTTTAAACCCACTTCAAAAACATGATGATGATCA 572
DB 587 TGTAAATGCTTTGGGAAATTTTAAACCCACTTCAAAAACATGATGATGATGATCA 646
QY 573 ACAA 578
DB 647 ACAA 652

RESULT 41
LOCUS CD723203 588 bp mRNA linear EST 26-JUN-2003
DEFINITION oj19b07.y1 Human lacrimal gland, unamplified: oj Homo sapiens cDNA
clone oj19b07 5', mRNA sequence.
ACCESSION CD723203
```

VERSION CD223203.1 GI:32274051
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 588)
 AUTHORS Dickinson, D., Laurie, G. and Mistow, G.
 TITLE Expressed sequence tag analysis of human lacrimal gland
 JOURNAL Unpublished (2002)
 COMMENT Contact: Mistow, G.
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: gdraeme@helix.nih.gov
 Plate: 19 row: b column: 07
 Seq primer: M13RPI reverse primer (ABI).
 Location/Qualifiers
 1..588
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="oj19b07"
 /tissue_type="lacrimal gland"
 /dev_stage="Adult"
 /lab_host="EMD10B"
 /note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted from 2 human lacrimal glands. A directionally cloned cDNA library in the pCMVSPORT6 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurie MD) essentially following the protocols of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor (5'-gactagttctagatgcgagcgccgccc(7)15-3'). EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."
 ORIGIN

Query Match 84.8%; Score 509.4; DB 6; Length 588;
 Best Local Similarity 99.8%; Pred. No. 3.2e-130;
 Matches 510; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 91 CTTGGGTAATAATGATGCAATTCATCGCCACACCTTAATAGAGACCAAGTAGTA 150
 DB 1 CTTGGGTAATAATGATGCAATTCATCGCCACACCTTAATAGAGACCAAGTAGTA 60
 QY 151 GGTGCTGATGAGCTCTACAGATGACCAAGGATACCAACCAAGGAGGAGGAG 210
 DB 61 GGTGCTGATGAGCTCTACAGATGACCAAGGATACCAACCAAGGAGGAGGAG 120
 QY 211 GAAGATCATCAAGAACTCTATCAAGACATCATCAAGCTGCGCATCTTTATAGAA 270
 DB 121 GAAGATCATCAAGAACTCTATCAAGACATCATCAAGCTGCGCATCTTTATAGAA 180
 QY 271 TCAGTTAATCAAGATGAGTACATTTGATGAGAAATTAAGAAAGTTCATAGCT 330
 DB 181 TCAGTTAATCAAGATGAGTACATTTGATGAGAAATTAAGAAAGTTCATAGCT 240
 QY 331 TCGTATGACCGGTGATGATTCATCAAGTATTAATCTTTGACCGAGATGTTATC 390
 DB 241 TCGTATGACCGGTGATGATTCATCAAGTATTAATCTTTGACCGAGATGTTATC 300
 QY 391 CAGGCTGTTAAAGATGAGAGATGCTGCACCAATCATTCAGGCGCACTCACTGC 450
 DB 301 CAGGCTGTTAAAGATGAGAGATGCTGCACCAATCATTCAGGCGCACTCACTGC 360
 QY 451 CAAGTACATGACGCGGTAAATGATCTTTGATCAATTTTTCAGATGTAATTTTGGC 510

Db 361 CAAGTACATGACGCGGTAAATGATGTTGATCAATTTTTCAGATGTAATTTTGGC 420
 QY 511 TCGCTGTATATATCTTTTGGGAATTTTAAACCCCACTTAACAAAATATGATGTAT 570
 DB 421 TCGCTGTATATATCTTTTGGGAATTTTAAACCCCACTTAACAAAATATGATGTAT 480
 QY 571 CAACAAATGTTGATGAGAGAAATATCA 601
 DB 481 CAACAAATGTTGATGAGAGAAATATCA 511
 RESULT 42
 BP345938
 LOCUS BP345938
 DEFINITION BP345938 Sugano cDNA library, synovial membrane Homo sapiens CDNA
 clone SYN08154, mRNA sequence.
 ACCESSION BP345938
 VERSION BP345938.1 GI:52275919
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 583)
 AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.
 Location/Qualifiers
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="SYN08154"
 /tissue_type="synovial membrane"
 /clone_lib="Sugano cDNA library, synovial membrane"
 ORIGIN

Query Match 82.8%; Score 497.8; DB 5; Length 583;
 Best Local Similarity 99.4%; Pred. No. 5.3e-127;
 Matches 510; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1 CTTACGCTCCCGGCGCCGCTGCG-CGACTCTCTCCATGCGCCACAGATGCTTTAATCC 59
 DB 71 CTTACGCTCCCGGCGCCGCTGCGCCACTCTCCGATGCGCCACAGATGCTTTAATCC 130
 QY 60 AAAACCTGGCCGTTTCAGGACCAAAAGAGATCTGGGTAAATGGGTCCAATTCATCG 119
 DB 131 AAAACCTGGCCGTTTCAGGACCAAAAGAGATCTGGGTAAATGGGTCCAATTCATCG 190
 QY 120 CCACCACTTAATAGACGACACAGTAGTAGGTGCTGATGAGCTCTACAGAGTACCA 179
 DB 191 CCACCACTTAATAGACGACACAGTAGTAGGTGCTGATGAGCTCTACAGAGTACCA 250
 QY 180 GGGAGTACACCCAAAACAAGAGAGGCGAGAGAAATCATCAAGAACTTCATTAAGAC 239
 DB 251 GGGAGTACACCCAAAACAAGAGAGGCGAGAGAAATCATCAAGAACTTCATTAAGAC 310
 QY 240 TCATCAAGCTGGCCATCTTTTATAGGAATTAATCAAGTATCAAGATGAGCTAGCTTGA 299
 DB 311 TCATCAAGCTGGCCATCTTTTATAGGAATTAATCAAGTATCAAGATGAGCTAGCTTGA 370
 QY 300 TGGAGAAATTTAAGAAAGTTCATCAAGCTTGTATGACCGGTGCTCAAGTTTCATCAAG 359
 DB 371 TGGAGAAATTTAAGAAAGTTCATCAAGCTTGTATGACCGGTGCTCAAGTTTCATCAAG 430
 QY 360 TGGATTAATCCTTTGACCGGAATGTTTATCCAGCGTGTAAATAGATCAAGAGATGC 419

```

Db      431 TGGATATACCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGAATGACAGAGATGC 490
Qy      420 TGCACCAATTCATTTGCGCCCACTCAGCCAGACATGACAGCGGTAAATATGCT 479
Db      491 TGCACCAATTCATTTGCGCCCACTCAGCCAGACATGACAGCGGTAAATATGCT 550
Qy      480 TGGATCATTTTCAGATGTGGAATTTTGGCTG 512
Db      551 TGGATCATTTTCAGATGTGGAATTTTGGCTG 583

RESULT 43
LOCUS   CR771295                605 bp    mRNA    linear    EST 23-SEP-2004
DEFINITION DKEP469M0735_r1_469 (synonym: pkid1) Pongo pygmaeus CDNA clone
VERSION   DKEP469M0735_5', mRNA sequence.
KEYWORDS  CR771295
SOURCE    CR771295.1 GI:52614568
ORGANISM  EST.
          Pongo pygmaeus (orangutan)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
          1 (bases 1 to 605)
AUTHORS   Pousetka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
          Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
          Wiemann,S.
          Pongo pygmaeus mRNA (Pousetka,A., Albert,R., Moosmayer,P., et al.)
          Unpublished (2004)
          CONTACT: MIPS
          MIPS
          Ingolstedter Landstr.1, D-85764 Neuberg, Germany
          This is the 5' sequence of the clone insert. Clone from S. Wiemann,
          Molecular Genome Analysis, German Cancer Research Center (DKFZ);
          Email s.wiemann@dkfz-heidelberg.de; any. Please contact R2PD for
          ordering:
          http://www.rzpd.de/cgi-bin/product/ci.cgi?cloneID=DKFZp469M0735
          Further information about the clone and the sequencing project is
          available at http://mips.gsf.de/projects/cdna/.
FEATURES
source    1..605
          /organism="Pongo pygmaeus"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="DKFZp469M0735"
          /tissue_type="kidney"
          /dev_stage="adult"
          /lab_host="DH10B"
          /clone_lib="469 (synonym: pkid1)"
          /note="Vector: pSPORT1_sfi; site_1: sfi1A; site_2: sfi1B"

ORIGIN
Query Match      82.6%; Score 496.6; DB 7; Length 605;
Best Local Similarity 99.2%; Pred. No. 1,2e-126;
Matches 499; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      36 TGGCCACAAATGCTTTAATTCGAAAACCTGGCCGTTCCAGGACAAAAGAAATCTTGG 95
Db      103 TGGCCACAAATGCTTTAATTCGAAAACCTGGCCGTTCCAGGACAAAAGAAATCTTGG 162
Qy      96 GTAATAATGTTGTCGAATTCATGCGACCACTTAATAGACACAAAGTAGAGAGTGC 155
Db      163 GTAATAATGTTGTCGAATTCATGCGACCACTTAATAGACACAAAGTAGAGAGTGC 222
Qy      156 TGGATGAGCTTACAGAGTACCGAGGAGTACCCAAAACAAAGAGAGGACAGAGAAG 215
Db      223 TGGATGAGCTTACAGAGTACCGAGGAGTACCCAAAACAAAGAGAGGACAGAGAAG 282
Qy      216 TCATCAAGAACTCCTCAAGACAGTATCAAGTGGCCATTTTATAGAAATATCACT 275
Db      283 TCATCAAGAACTCCTCAAGACAGTATCAAGTGGCCATTTTATAGAAATATCACT 342

```

```

Qy      276 TTAATCAAGATAGCTAGCATTTGATGAGAAATTTAAGAGAAAGCTTCAGCTGTCTA 335
Db      343 TTAATCAAGATAGCTAGCATTTGATGAGAAATTTAAGAGAAAGCTTCAGCTGTCTA 402
Qy      336 TGAACGCTGTCAGTTTCCATCAGTGTGATTTATACCTTGGACCGGAATGTTATCCAGGC 395
Db      403 TGAACGCTGTCAGTTTCCATCAGTGTGATTTATACCTTGGACCGGAATGTTATCCAGGC 462
Qy      396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCACTGCCAAGT 455
Db      463 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCACTGCCAAGT 522
Qy      456 CACATGACCGGTTAATATGCTTTGATCATTTTTCAGATTTGTAATTTTGGTGCCT 515
Db      523 CACATGACCGGTTAATATGCTTTGATCATTTTTCAGATTTGTAATTTTGGTGCCT 582
Qy      516 TGTATATCTTTTGGCAATTTT 538
Db      583 TGTATATCTTTTGGCAATTTT 605

RESULT 44
LOCUS   BI760235                744 bp    mRNA    linear    EST 25-SEP-2001
DEFINITION 603045171P1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5185720 5',
          mRNA sequence.
VERSION   BI760235
KEYWORDS  BI760235.1 GI:15751813
SOURCE    EST.
          Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 744)
AUTHORS   NIH-MGC http://mgs.mci.nih.gov/.
          NIH-MGC http://mgs.mci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          CONTACT: Robert Strausberg, Ph.D.
          Email: cga@bbs-remail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/ILNL at:
          http://image.llnl.gov
          Plate: L1AM11463 row: m column: 17
          High quality sequence stop: 743.
FEATURES
source    1..744
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="IMAGE:5185720"
          /lab_host="DH10B"
          /clone_lib="NIH MGC 116"
          /note="Organ: pooled colon, kidney, stomach; Vector:
          pCMV-SPORT6; site_1: NotI; site_2: EcoRV (destroyed); RNA
          source anonymous pool of 3 colons, age 26 yo male, 49 yo
          female, 71 yo male colon; 46 yo male kidney, and pool of 2
          stomachs, 62 yo male and 70 yo female. Library is
          oligo-dT primed and directionally cloned (EcoRV site is
          destroyed upon cloning). Average insert size 1.4 kb,
          insert size range 1-3 kb. Library is normalized and
          enriched for full-length clones and was constructed by C.
          Gruber (Invitrogen). Research Genetics tracking code
          023. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      82.3%; Score 494.6; DB 4; Length 744;
Best Local Similarity 97.5%; Pred. No. 4,4e-126;
Matches 555; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

```

QY 36 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTGCAGGACAAAGAGATCTGG 95
 |||||
 DB 175 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTGCAGGACAAAGAGATCTGG 234
 |||||
 QY 96 GTAAATGCTGTCCTTAAATCCATGCGCCACCACTTAAATAGACACACAAGTAGTAGGTCG 155
 |||||
 DB 235 GTAAATGCTGTCCTTAAATCCATGCGCCACCACTTAAATAGACACACAAGTAGTAGGTCG 234
 |||||
 QY 156 TGGATGAGCTTACAGAGTAGACCAAGGAGTAGACCCCAAAACAAAGAGAGGACAGAGAGA 215
 |||||
 DB 295 TGGATGAGCTTACAGAGTAGACCAAGGAGTAGACCCCAAAACAAAGAGAGGACAGAGAGA 354
 |||||
 QY 216 TCATCAAGAACCTTCATCAAGACAGTCAATGAGTGCCTTTATAGGAATATCACT 275
 |||||
 DB 355 TCATCAAGAACCTTCATCAAGACAGTCAATGAGTGCCTTTATAGGAATATCACT 414
 |||||
 QY 276 TTAATCAAGATGAGTCAATGATGAGAAATTTAA-GAAGAAAGTTCATCAGTTGCT 334
 |||||
 DB 415 TTAATCAAGATGAGTCAATGATGAGAAATTTAA-GAAGAAAGTTCATCAGTTGCT 474
 |||||
 QY 335 ATGACCGTGGTCACTTTCATCAAGTGTATTAATCTTTGACCGGAATGTTATCCAGG 394
 |||||
 DB 475 ATGACCGTGGTCACTTTCATCAAGTGTATTAATCTTTGACCGGAATGTTATCCAGG 534
 |||||
 QY 395 CTGTTAAATGAATGCA-GAGAGATGCTGACCAATATTCAGGCGCACTCACTGCCAA 453
 |||||
 DB 535 CTGTTAAATGAATGCA-GAGAGATGCTGACCAATATTCAGGCGCACTCACTGCCAA 594
 |||||
 QY 454 GTCACTGAGCGGGTTAATATGCTTGTATCATTTTTCAGTTGTGATTTTGGCTG 513
 |||||
 DB 595 GTCACTGAGCGGGTTAATATGCTTGTATCATTTTTCAGTTGTGATTTTGGCTG 652
 |||||
 QY 514 CTGTATTAATCTTTTGGAAATTTTAAACCCCACTTAAACAAAC-TATGATGATGATCA 572
 |||||
 DB 653 CTGTATTAATCTTTTGGAAATTTTAAACCCCACTTAAACAAAC-TATGATGATGATCA 712
 |||||
 QY 573 ACAAAATGTTGATGAAGAGAACATATGA 601
 |||||
 DB 713 ACAAAATGTTGATGAAGAGAACATATGA 741
 |||||

RESULT 45
 BP433798 773 bp mRNA linear EST 30-DEC-2003
 LOCUS BP433798 full-length enriched swine cDNA library, adult lung Sus
 DEFINITION scrofa cDNA clone UNG010079A10 5', mRNA sequence.
 ACCESSION BP433798
 VERSION BP433798.1 GI:40423865
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 773)
 Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,
 Okumura,N., Hamasaka,N. and Awata,T.
 PEDR (Pig EST Data Explorer): construction of a database for ESTs
 derived from porcine full-length cDNA libraries
 Nucleic Acids Res. 32 (1), D484-D488 (2004)
 Contact: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in
 Animal Genome Research Program (Japan) by National Institute of
 Agrobiological Sciences and STAFF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA
 library
 Vector sequences were eliminated by RepeatMasker version 2002/07/13
 and crossmatch version 0.990319

FEATURES
 Source
 1.773
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="UNG010079A10"
 /issue_type="adult"
 /dev_stage="adult"
 /clone_11b="full-length enriched swine cDNA library, adult
 lung"

ORIGIN
 Query Match 81.4%; Score 489.2; DB 5; Length 773;
 Best Local Similarity 91.5%; Pred. No. 1.4e-124;
 Matches 518; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 36 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTGCAGGACAAAGAGATCTGG 95
 |||||
 DB 155 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTGCAGGACAAAGAGATCTGG 214
 |||||
 QY 96 GTAAATGCTGTCCTTAAATCCATGCGCCACCACTTAAATAGACACACAAGTAGTAGGTCG 155
 |||||
 DB 215 GTAAATGCTGTCCTTAAATCCATGCGCCACCACTTAAATAGACACACAAGTAGTAGGTCG 274
 |||||
 QY 156 TGGATGAGCTTACAGAGTAGACCAAGGAGTAGACCCCAAAACAAAGAGAGGACAGAGAGA 215
 |||||
 DB 275 TGGATGAGCTTACAGAGTAGACCAAGGAGTAGACCCCAAAACAAAGAGAGGACAGAGAGA 334
 |||||
 QY 216 TCATCAAGAACCTTCATCAAGACAGTCAATGAGTGCCTTTATAGGAATATCACT 275
 |||||
 DB 335 TCATCAAGAACCTTCATCAAGACAGTCAATGAGTGCCTTTATAGGAATATCACT 394
 |||||
 QY 276 TTAATCAAGATGAGTCAATGATGAGAAATTTAAAGAAAGTTCATCAGTTGCT 335
 |||||
 DB 395 TTAATCAAGATGAGTCAATGATGAGAAATTTAAAGAAAGTTCATCAGTTGCT 454
 |||||
 QY 336 TGAACCGTGGTCACTTTCATCAAGTGTATTAATCTTTGACCGGAATGTTATCCAGG 395
 |||||
 DB 455 TGAACCGTGGTCACTTTCATCAAGTGTATTAATCTTTGACCGGAATGTTATCCAGG 514
 |||||
 QY 396 TGTAAATGAATGCAAGAGATGCTGACCAATATCTTAAAGGCGCACTCACTGCCAAGT 455
 |||||
 DB 515 TGTCAATGATGCAAGAGATGCTGCTCCACAGATCATCAAGGCTCACTCAAGGCCAAGT 574
 |||||
 QY 456 CACATGAGCGGGTTAATATGCTTGTATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
 |||||
 DB 575 CACATGAGCGGGTTAATATGCTTGTATCATTTTTCAGATTGTGAATTTTGGCTGCT 634
 |||||
 QY 516 TGTATTAATCTTTTGGAAATTTTAAACCCCACTTAAACAAACATGATGATGATCAACA 575
 |||||
 DB 635 TGTATTAATCTTTTGGAAATTTTAAACCCCACTTAAACAAACATGATGATGATCAACA 694
 |||||
 QY 576 AAATGTTGATGAAGAGAACATATGA 601
 |||||
 DB 695 AAATGTTGATGAAGAGAACATATGA 720
 |||||

Search completed: July 28, 2005, 19:01:12
 Job time : 2905 secs

THIS PAGE BLANK (USPTO)